

כח

RESULT 8				
LOCUS	ATU41339			
DEFINITION	Arabidopsis thaliana ANT	1905 bp	linear	PLN 23-OCT-1996
ACCESSION	U41339			
VERSION	U41339.1	GI:1244707		
KEYWORDS				
SOURCE	Arabidopsis thaliana			
ORGANISM	Arabidopsis thaliana			

Wed Mar 17 08:18:55 2004

us-10-024-632-2.rgt

Page 13

Db 581 GGA----- 583  
Qy 218 GInGInMetAsnCyeGlyMetGlyangGluArgangGlyValSerLeuGlySerValGly 237  
Db 584 GGGCTCTTTAATGATGAGGGATATATGAGGAATTCA----- 619  
Qy 238 CyeGlyGlyLeuGInSerLeuSerLeuSerMetSerProGlySerGInSerSerCyeVal 257  
Db 620 -----CAGTCACTGAGCTTATCATGAGCCCGGCTCAGATCTAGCTGCATC 667  
Qy 258 ThrAlaProSer----- 261  
Db 668 ACTGGCTCTTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 727  
Qy 262 -----GlyThrAspSerValAlaVal 268  
Db 728 CAGATCTGAGCTCTGTCGACAGACAGCTGGGTTGAGACACACAGATGGCGCT 787  
Qy 269 AspAlaLeuLeuArgGlyHis-----AlaLeuLeuGlyGlnGlyGlnProValHis 285  
Db 788 GCGAAG 847  
Qy 286 ArgLeuSerLeuAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
Db 848 AGAAATCTATCATCTTTCGACACAGACTTCCTCATTCAGAGCGCTTTCAGAGACT 907  
Qy 306 ArgGlyThrGlyArgGlyGlnAlaHisLeuTyrAspAsnSerCysLeuLeuGlyGln 325  
Db 908 AGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967  
Qy 326 ThrArgGlySerGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlyLeuAlaArg 345  
Db 968 AGTGAAG 1027  
Qy 346 AlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHisLeuAsnProSerLe 365  
Db 1028 GCATATATCTTCTGCTGCTCAAGTACTGGGGTCCCTCTACCTCACACCAATTTCTGCG 1087  
Qy 366 GluAsnTyrGlnValGlnLeuGlnGlnMetLeuAsnMetSerArgGlnGlnTyrValAla 385  
Db 1088 GAGAAATTCAGAAAGAGATTCAGAAAGATTCAGAAAGATTCAGAAAGATTCAGAAAGAT 1147  
Qy 386 HisLeuArgArgLeuSerSerGlyPheSerArgGlyAlaSerLeuTyrArgGlyValThr 405  
Db 1148 CATTTGAG 1207  
Qy 406 ArgHisGlnHisGlyArgTyrGlnAlaArgLeuArgValAlaGlyValAsnLysAsp 425  
Db 1208 AGACATCACAGCATGAG 1267  
Qy 426 LeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaGlnAlaTyrAspValAlaAla 445  
Db 1268 CTCACTGAGCTTTCGACCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1327  
Qy 446 IleuysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGln 465  
Db 1328 ATTAAGTTCCGTCGACAAATGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 1387  
Qy 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGlnLeuAlaArgArgLysLysAsnAsn 485  
Db 1388 CGTATCATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1438  
Qy 486 AspProArgAsnLysAspIleAspTyrAsnLysSerValTyrSerValAsnGln 505  
Db 1439 -----ACACACAGCATGCTC-----GTCCAGAAATACT 1465  
Qy 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGluTyrLysMet 525  
Db 1466 GAA----- 1468  
Qy 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
Db 1469 -----GACCAAAACC 1477

Qy 546 IleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
Db 1478 GCTCTAAATGCT----- 1489  
Qy 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerLysIleGly 585  
Db 1489 ----- 1489  
Qy 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGlnAlaSer 605  
Db 1490 -----GTTGTGAAGAGCTGCTTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1522  
Qy 606 ProGlnLysArgGlyProSerLeu-----LeuPheProMetProProMetGlnThrLys 623  
Db 1523 ACTCCGAGAGACTCTGAGTTTCCGCGAGTTTCCGCGAGTTTCCGCGAGTTTCCGCGAGTT 1582  
Qy 624 IleVal-----AsnProIleGlyThrSerValThrSerThrPheProSerProThrValGln 642  
Db 1583 ATGTTGAGATCAAAATATGAGGCGAGAAATATGAGCTTTCGACATCAACCTAAATGCTGAG 1642  
Qy 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrPThrAsp 662  
Db 1643 CTTAAG-----ACCGGCGCTCTTACTTTCCTCAGATGCCGCTTTCGCTGCTGAGT 1699  
Qy 663 Thr 663  
Db 1700 TCT 1702  
PRTTCT 9

Result No.	Score	Query Match	Length	DB	ID	Description
1	1328	37.9	555	2	S71365	ovule development
2	922.5	26.4	425	2	B96750	hypothetical prot
3	916.5	26.2	581	2	T51580	ovule development
4	890	25.4	485	2	T03638	hypothetical prot
5	879	25.1	516	2	F96549	hypothetical prot
6	831	23.7	566	2	T49988	ovule development
7	645	18.4	332	2	D86295	hypothetical prot
8	616	17.6	308	2	H96827	protein F20H17.12
9	534	15.2	485	2	S54116	hypothetical prot
10	517	14.8	205	2	T47591	antigenmueta-like
11	517	14.8	432	2	B84336	ABE71A2 protein
12	497.5	14.2	436	2	B84845	probable Ap2 doma
13	486.5	13.9	464	2	C84686	probable Ap2 doma
14	483	13.8	446	2	T03981	ABE71A2-like pro
15	467.5	13.3	433	2	T03574	intermediate spli
16	286.5	8.2	247	2	T06721	hypothetical prot
17	275	7.9	236	2	T02577	probable Ap2 doma
18	161.5	4.6	262	2	E96747	hypothetical prot
19	140	4.6	314	2	E96492	protein P505.5
20	158.5	4.5	245	2	B86166	hypothetical prot
21	158	4.5	292	2	T03030	hypothetical prot
22	157.5	4.5	199	2	B86288	F501.31 protein
23	156.5	4.5	248	2	T07728	transcription fac
24	155	4.4	358	2	D96579	hypothetical prot
25	152.5	4.4	336	2	E84594	Ap2 domain trans
26	151.5	4.3	268	2	T04787	hypothetical prot
27	151.5	4.3	334	2	T02896	hypothetical prot
28	150	4.3	256	2	H96837	unknown protein T
29	150	4.3	295	2	T00399	probable Ap2 doma

30	147.5	4.2	343	2	T05607	hypothetical prote
31	147.5	4.2	1475	2	P86399	protein F47L21.22
32	146	4.2	303	2	T04541	hypothetical prote
33	144	4.1	212	2	T49580	hypothetical prote
34	144	4.1	225	2	T02433	DNA binding protei
35	143.5	4.1	272	2	T05015	hypothetical protei
36	143.5	4.1	272	2	T03927	DNA binding protei
37	141.5	4.0	344	2	T53329	DNA binding protei
38	141	4.0	236	2	B84732	hypothetical prote
39	140.5	4.0	171	2	T00432	probable AP2 domai
40	140	4.0	283	2	C84828	AP2 domain transcr
41	139.5	4.0	222	2	T53019	ethylene responsiv
42	139	4.0	218	2	T01076	transcription fact
43	139	4.0	352	2	T51330	DNA binding protei
44	137.5	3.9	204	2	B86410	protein F3M18.20 [
45	137.5	3.9	328	2	G84826	hypothetical prote

## ALIGNMENTS

RESULT 1  
S71365  
ovule development protein a1nregcontaining - Arabidopsis thaliana  
N.alternate names: A42 domain-containing protein; proteoin T28119..30  
C.species: Arabidopsis thaliana (mouse-ear cress)  
C.date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 22-Oct-1999  
C.accession: S71365; T06022  
R.Klucher, K.  
submitted to the EMBL Data Library, November 1995  
A.reference number: S71365  
A.accession: S71365  
A.molecule type: mRNA  
A.residues: 1-555 <KLU>  
A.cross-references: EMBL:U04256; NID:G1209096; PID:G1209099  
R.Bevan, M.; Van Der Schueren, J.; Chang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.  
submitted to the Protein Sequence Database, March 1999  
A.reference number: Z15484  
A.accession: T06022  
A.molecule type: DNA  
A.residues: 1-555 <BEV>  
A.cross-references: EMBL:AL035709; GSPDB:GN00066; ATSP:T28119..30  
A.experimental source: cultivar Columbia; BAC clone T28119  
C.genetics:  
A.gene: ATSP:T28119.30; ANT  
A.map position: 4  
A.introns: 111/1; 289/2; 317/1; 320/1; 349/3; 391/2; 417/1

Query Match	Score	DB 2;	Length
37.9%	1328;	DB 2;	555

Best Local Similarity 43.5%; Pred. NO. 5.1e-86;  
Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22

QY 1 MKRINSSNNTDGNHNLGFSLSPH-MKY-----EATSATVPTTFYM 43  
.....  
.....

Db 1 MKSFCNDNDNNHNTNTNLGFSLSNNMMKMGRCGRATYSSSTSSAATSSSSVPPQLV 60

QY 44 SPSQSHJNFGNCYGVGENGNFHSPLTVNPLKSDGLCIIEALKRSGTQVMVFTSSPKIE 103

Db 61 G---DNTSNFGVCYGSNPNGGITSHMSVNPPLRSDDSLCLMEALNRRSSHSHHHQDSSPKYE 117

```
Qy      104 DFLGATMGTHGYSGHERG--LSLDSIYNSQNAAQPNRLLSQPF-----HQGGHMSVQ 157
```

Db 118 DFFG----THHNTSHKAMDLSLDFNTTH----EPNTTNFQEEFSFPQYRNH---- 166

167	PREP- DNYCND-----	DSTYTC-----	183
QY	158 THPYISGLACHGYDAPUSEEITNEIHVDCSSLMFOIMBSJANWVAF IRETS INCVOL	21	

```
DB -----EDEL-RUNSWD-----FOUJING----- 260
167 -----
C: 318 COMMCGMCPBNCVSTCSWACGCEI--OSTSLSMSBQSOSSCYTAPB 261
```

Dh  
183 -----GSENNVY-----YGEFOOSLSMSPGSOSSCTTGSHHHQQNQNHNHOSQNH 229

00 262 -----GTDSVAVDAKRGH---AKLGOKOPVHRKSIDTFGORTSQYRGVTR 304

```

Db      230 QQISEALVETSVGFETTTWAAAKKKGGQGDVYVVGQKQVHRSKIDTGGQTSQVRGTR 289
Qy      305 HEMTGRYEALWDNSCKEKGQTRKGRQVYLGQYDMEKRAAYDLAALKWGPSTHIFS 364
Db      290 HEMTGRYEALWDNSCKEKGSRKGRQVYLGQYDMEKRAAYDLAALKWGPSTHIFS 349
Qy      365 IENYQVQLEEMKMSQEVAAHLRRKSGFSRGASITGVTRHHQHGMRARIGRVANR 424
Db      350 AENYQVQLEEDMKMTQEVAAHLRRKSGFSRGASITGVTRHHQHGMRARIGRVANR 409
Qy      425 DYLGTFTSTQEEAAEYDVAALKFRGANAVTNFDISRVDYERIMASNLGELARRKD 484
Db      410 DYLGTFTSTQEEAAEYDVAALKFRGTNAVTFNFDITRVDYERIMASNLGELARRN 467
Qy      485 NDRPRNDIDYNSVYTVNNEETVQVQAGNNNNDESEKMYLFPHPSQQQANNGSDQ 544
Db      468 -----NNSTIV-VKNTT-----DQ 479
Qy      545 KIMNCGYNSAFSMALODLIGDSVSGQHMLDESSKIGTHFNTSLVTSLSRPA 604
Db      480 TALNA-----VEGGSNREV 494
Qy      605 SPEKRPRL-LPRMPMETKIV-NPIGTSVTSMLPSPTVQMRPSPALSLSLPFAWT 661
Db      495 STPERLLSPFAIFALPQVQKQFSGNMGNMSPMTSNPNAELK-TVALTLPPQVFAAMA 553
Qy      662 DT 663
Db      554 DS 555

RESULT 2
B96750
hypothetical protein F28P22.24 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 31-Mar-2001
C:Accession: B96750
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Comai, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; NUID:21016719; PMID:11130712
A:Accession: B96750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <STO>
A:Cross-references: GB:A005173; NID:96648171; PIDN:AAF21171.1; GSPDB:GMO0141
C:Genetics:
A:Gene: F28P22.24
A:Map position: 1

Query Match      26.4%; Score 923.5; DB 2; Length 425;
Best Local Similarity 43.4%; Pred. No. 1.5e-57;
Matches 214; Conservative 60; Mismatches 108; Indels 111; Gaps 16;

```

```

Qy      192 MPQMTES-----LKNVV---APTRRSTHQVLEQNMCGM---GNERNGVSL 233
Db      110 DPNLSNDYGGFEFVSGSVSFKSLMLEQTPAFLPSHSYVTEACTSNNHSPNEEFGYNT 169
Qy      234 -GSGVGGELQSLSLSSPGSSQSVTAPSGT---DSVAVDK-KRGAHLGQKQPVHR 287
Db      170 NGSM-----LSLALSHGACDLINESNVSARVEEPVVKDEKRKLLVYPQYKESVPRK 222
Qy      288 SITFGQRTQYGVTRHRRTGTYEALWDNSCKEKGQTRKGRQVYLGQYDMEKRAAY 347
Db      223 SVDSYGGRTSYGVTRHRRTGTYEALWDNSCKEKGQTRKGRQVYLGQYDMEKRAAY 282
Qy      348 DLALKWGPSTHIFNSIENYQVQLEEMKMSQEVAAHLRRKSGFSRGASITGVTRH 407
Db      283 DLALKWGPSTHIFNPLSYKEKEIEDLNMMNQEFVAMLRNSGFSRGASITGVTRH 342
Qy      408 HQGRWQARIGRVANRDLTLGTFTSTQEEAAEYDVAALKFRGANAVTNFDISRVDYER 467
Db      343 HQGRWQARIGRVANRDLTLGTFTSTQEEAAEYDVAALKFRGANAVTNFDINRYDKRI 402
Qy      468 MASSNLGELAR 480
Db      403 CSSSTIVDSQAK 415

RESULT 3
T51580
coulie development protein atnegenuta-like protein - Arabidopsis thaliana
N:Alternate names: protein T10B6_90
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #ext_change 18-Aug-2000
C:Accession: T51580
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Aaamizu, E.; Kotani, H.; Tabata, S.; M
submitted to the Protein Sequence Database, August 2000
A:Reference number: 225394
A:Accession: T51580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-581 <SAT>
A:Cross-references: EMBL:AJ391142
A:Experimental source: cultivar Columbia; BAC clone T10B6
C:Genetics:
A:Map position: 5
A:Introns: 78/1; 216/2; 244/1; 273/3; 298/2; 315/2; 341/1
A:Note: T10B6_90

Query Match      26.2%; Score 916.5; DB 2; Length 581;
Best Local Similarity 35.2%; Pred. No. 7.5e-57;
Matches 255; Conservative 76; Mismatches 170; Indels 223; Gaps 25;

```



Db 202 ESPGQRTSIYRGVTRHRTGRYEAHLMDNSCKREGQTRKGRQ---GGYDKEEKAARAYDL 258

QY 350 AALKXWGPSTHINFSLENNVOOLEEMKMSRQEVYAHLRKSSGFSRNASIYRQVTHHQ 409

Db 259 AALKXWGTITTTNPFSEYEKEVEEMKMTROEYVARSRSRSGFSRNASIYRQVTHHQ 318

QY 410 HGRWQARIGRVAGNCDLYLGTFTSTDEAAAYDVAAIKFRGANAVTNFDSRYDVERIMA 469

Db 319 HGRWQARIGRVAGNCDLYLGTFTSTDEAAAYDVAAIKFRGANAVTNFDMRYNKKALE 378

QY 470 SSNLLAGELARRKNDPKNKIDYKSVYTVNNEETVOQAGN---NNNENDEMKVYL 527

Db 379 SSELPLTGSSAKRLKD-----VNNPVPAMMTSNVSSANNVSGWQMTA 421

QY 528 FNHPS-----OQOQANGSDQKIMCNGYRNBAFSAALQDLIGDVSQGHMLDE 580

Db 422 FOHQCMQDLSLQOQOERYVG---YNGGNLSTESTRCFK-----QEEEOQHFLNRS 471

QY 581 SSKIGT--HRSNNSLVTSLSSRSRSPKRGPSLLFMPMPETKIVNPIGTSTVMSLPS 638

Db 472 PSMHTVVDHSSITSDSVTVCGVVSYGCGFPAL-----PVGTSV-NYDPF 517

QY 639 PTV-----QMRSP-----AISLSH-----LPVYAS 659

Db 518 TAAEIVANARNHYAAHQOQOQOIQSPGDFPVALSNHSSNMYFHGEGGEGAPFTSV 577

QY 660 WTD 663

Db 578 WMDT 581

## RESULT 4

T03638  
hypothetical protein - maize  
C/Species: Zea mays (maize)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 20-Jun-2000  
C/Accession: T03638  
R/Daniel, T.J.  
submitted to the EMBL Data Library, January 1995  
A/Reference number: Z14985  
A/Accession: T03638  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-485 <DAN>  
A/Cross-references: EMBL:Z47554; NID:g1103747; PTDN:CAA87634.1  
A/Experimental source: strain Black Mexican Sweet

Query Match 25.4%; Score 890; DB 2; Length 485;  
Best Local Similarity 44.2%; Pred. No. 4.2e-55;  
Matches 200; Conservative 45; Mismatches 94; Indels 114; Gaps 10;

QY 187 DCSLMPQMTGKXNVAPTRFSTHQVLEQNMCG--MGENNGIS-----LG 234

Db 4 DMSAYPH-----HMLFSLSNNYHGLLEAFSSSGTPTGDEGAVESSPRVEDPLG 57

QY 235 SVG-----CGELQSLSLSMSFGSOSG-----CYTASGT 263

Db 58 GVCVCGAPRSRLQIRITSLCAASCSITARFLRYPAAGSGTTVGBPLSFTLAASST 117

QY 264 DSAVAVAKKRGAAKQKQKQVARKSIDTQRTSYRGVTRHRTGRYEAHLMDNSCKE 323

Db 118 DVMAASDDQS-----RSAETRGQRTSISYRGVTRHRTGRYEAHLMDNSCKE 165

QY 324 GQTRKRGQVYLGQYDMEKARAYDLAALKXWGPSTHINFSLENNVOOLEEMKMSRQEV 383

Db 166 GQTRKRGQVYLGQYDMEKARAYDLAALKXWGPSTHINFSLENNVOOLEEMKMSRQEV 225

QY 384 VAHLRRSGSFGSISYGVGRHGHGWOARIGRVAGNCDLYLGTSTDEAAAYDV 443

Db 226 IASLRKSSGFSRNASIYRQVTHHQHGRWQARIGSVAGNCDLYLGTSTDEAAAYDV 285

QY 444 AAIKFRGANAVTNFDSRYDVERIM-----ASSNLLAG 476

Db 286 AAIKFRGANAVTNFDMRYNKKALE-----DYKSVYTVNNEETVOQAGN 515

QY 477 ELARRKNDPKNKID-----DYKSVYTVNNEETVOQAGN 515

Db 346 AAAAQATMPPEKQKYSWLLALHQQOQOQERQFPASAEYAYGGGVVDPMTGTSNGN 405

QY 516 NNNENDEMKVYL-----KXVLEFNHPSQOQOQANGSD 543

Db 406 NNTGSGVWVGATGSAVVGQOQSSSKQNGYASN 438

## RESULT 5

P96549  
hypothetical protein F11M5.6 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: P96549  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K  
amen, N.R.; Hughes, B.; Huizar, L.  
Nucleotide 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Ventner, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A06141; MUID:21016719; PMID:11130712  
A/Accession: P96549  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-516 <STO>  
A/Cross-references: GB:A0605173; NID:G4636931; PTDN:AAD30633.1; GSPDB:GN00141  
A/Genes: F11M5.6  
A/Map position: 1

Query Match 25.1%; Score 879; DB 2; Length 516;  
Best Local Similarity 38.4%; Pred. No. 2.8e-54;  
Matches 217; Conservative 70; Mismatches 140; Indels 138; Gaps 17;

QY 100 PLEDFLIGATMGTEYVSHRGLSLDSIYVNSQNAARPRDLSPFPQOQGMSTVQTH 159

Db 15 PKYADFLGYSKSGDH-----HTDNLV 36

QY 160 PYSGLACHGYQAPLEBETTKETHVSDCSLMPQMTGKXNVAPTRFSTHQVLEQ 219

Db 37 PY-----NDIHQ-----TNASDYFFQTNLPLTVTCASN--AP-----NNYELQESA 77

QY 220 KMGCGNERNVSGVSGVCGELQSLSLMSFP-----GSQSCVTAFPSTGDSVAVDAKRGH 275

Db 78 HNL-----OSLTSMSGSTAAALAAVATYKASAEISADNSSTTVTSGALV----- 124

QY 276 AVLQKQKQVHRSKSIDTQRTSYRGVTRHRTGRYEAHLMDNSCKEKGQTRKRGQVYLG 335

Db 125 -----EATPRRTLETFQRTSISYRGVTRHRTGRYEAHLMDNSCKEKGQTRKRGQVYLG 175

QY 336 GYDMEKARAYDLAALKXWGPSTHINFSLENNVOOLEEMKMSRQEVYAHLRKSSGFS 395

Db 176 GYDMEKARAYDLAALKXWGPSTHINFSLENNVOOLEEMKMSRQEVYAHLRKSSGFS 235

QY 396 RGASISYRGVTRHGHGWOARIGRVAGNCDLYLGTSTDEAAAYDVAAIKFRGANAVT 455

Db 236 RASAVYRGVTHHGHGWOARIGRVAGNCDLYLGTSTDEAAAYDVAAIKFRGANAVT 295

QY 456 NFDISRYDVERIMASNN--LAGELARRKNDPKNKIDYKSVYTVNNEETVOQAGN 515

Db 296 NFEINRYDKALIBENLPIGGAARKKEQALLESRRKKEEMTALDSNTHQYGAAGSS 355

QY 501 SYNNEETVOQ-----AGNNNE-----NDESEKMYLFNHPGQOQ 536

Db 356 SVASSSRLOQPPYPLSLQPPFHLHHQPLTLQNNNDISQYHDSFSYIOTQHLHQOQT 415

QY 537 ANNGSPQKIMNCGNRNSAFSWALDDLGIGSVSGGHNMDSESKGTGFSNYSLLVT 598  
 Db 416 NN-----YLGSSHTSQLYNNAYLGNSPGL-----LHGFSVDNNNTSGLGNNNGIG 463  
 QY 597 SLSSSRASPEKRPGL--FMP 619  
 Db 463 S-STVGSASAEFFPAKYDYMP 486

## RESULT 6

ovule development protein-like - Arabidopsis thaliana  
 N:Alternate names: protein Fl2B17.140  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #ext\_change 02-Jun-2000  
 C:Accession: I439388  
 R:Byatn, M.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z25026  
 A:Accession: I439388  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-566 <REV>  
 A:Cross-references: EMBL:AL353955, GSPDB:GN00063; ATSP:Fl2B17.140  
 C:Experimental source: cultivar Columbia; BAC clone Fl2B17  
 C:Genetics:  
 A:Gene: ATSP:Fl2B17.140  
 A:Map position: 5  
 A:Introns: 33/1, 259/2, 287/1, 316/3, 341/2, 384/1

Query Match	23.7%;	Score 831;	DB 2;	Length 566;
Best Local Similarity	36.1%;	Pred. No. 8e-51;		
Matches 223; Conservative	69;	Mismatches 167;	Indels 158;	Gaps 20;

```

QY 1 NNNHWTJGFSLSPPHKREATSA-----TWPTMYMSQOCHLSNFMQCYGVANGNFH 66
Db 30 NPYMSQKQKQKQEFFKEEACUAAAASADTILTFVDP-QSH-----70
QY 67 SPLTWMLKSDGSLCTILEALKRSQTOVMPTSSPKLEDFLGATNGTHEYSHBERGLSLD 128
Db 71 -----HSONHI-----PKLEDFLGDS-----87
QY 127 SIYYSNQAENAEPNRDLISOPFROOAHMVOT-----HPYSGLACHLYOAPLEETTK 181
Db 88 SIYRSPNOSTQPDQDSLSITQYDPRHHN-QGFSYSDHDFWTMA--GFOA---FSTNS 141
QY 182 ETVHSDCSSLMPOMTBGLKNWVAFTREBETHQOVLQOMNCMGENRNGVSLSGCEL 241
Db 142 GSEVDDSAISGRTHLAG-----DYLGHVV-----ESSGPELGHG-GST 179
QY 242 QSLISMSFSSQSS-----CTAPSGTDSVAVDKKGKHAFLKQXQPVH 285
Db 180 GALSISYNNVNNNNHNDNDNHVRGNNNERINNNNNDNEXTBSEKKAVALVETDSCS 239
QY 286 RKS1-DTFFQRTQYGVTRHRTGTYEALMDNSCKKGQTRKRGQVYLTGDMEEKAA 344
Db 240 NKKIADTFGGRTIYGVTRHRTGTYEALMDNSCRREGQARKRCQ--GGYKEDKAA 296
QY 345 RAYDLALKWYGSTHINFSIENYQVLEEMKMSROEVAHLRKSXSGFSRGSATYRGV 404
Db 297 RAYDLALKWMAATATTPTTNYSEVEMGMHTKOEIASLRKXSGFSRGSATIRGV 356
QY 405 TRHHQGRWQARIGRAYGNKDLYLGTFSQCEAAALYVAAIKFPGANAATNPISRYDV 464
Db 357 TRHHQGRWQARIGRAYGNKDLYLGTFAEEBAALYVAAIKFPGIAVYNFEMNRYDV 416
QY 465 ERIMASNTLAGELARKKQND-----PRKKDIDYKSVYTSYNNMETVQVQAGANNEND 520
Db 417 EALNKSA1PFGAARFKLKLSEAAASBQKPLIHGHQJHHFQOQOQOQOQLQSSPWHSS 476
QY 521 SEWKM-----VLFNHSQOQOQANGNSDQKIMNCAYNKAFAFM 555
Db 477 INFPLCNSAVSQIIPCGIPREAAALVHHHQOQOQOQOQOQOQNFQO--HFPANAS- 533

```

```
QY 560 ALQDLIGIDSVGSGQH 576
      ||| ::
Db 534 -----DSTGSNNNS 542
```

## RESULT 7

hyperchemical protein T2AD18.16 - Arabidopsis thaliana  
 CSpecies: Arabidopsis thaliana (mouse-ear cross)  
 CDate: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 CAccession: D86295  
 R|Theliosius, A.; Ecker, U. R.; Palm, C. U.; Federspiel, N. A.; Kaul, S.; White, C.; Alonso  
 Chin, C. W.; Chung, M. K.; Conn, L.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar, K.  
 ansen, N. F.; Hughes, B.; Hutzler, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C. A.; Li, J.; Rooney, T.; Rowley, D.; Saku, S. X.; Liu, Z. A.; Luoro, J. S.; Maiti, R.; Marziani  
 Rizzo, M. R.; Rooney, T.; Rowley, D.; Saku, S. X.; Liu, Z. A.; Luoro, J. S.; Maiti, R.; Marziani  
 A:Authors: Salzberg, S. L.; Schwartz, J. R.; Shinn, P.; Southwick, A. M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C. M.; Venter, J. C.; Davis, R. W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; PMID:21016719; PMID:11130712  
 A:Accession: D86295  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-333 <STO>  
 A:Cross-references: GB:AE005172; NID:g6587812; P|DN:AA18503.1; GSPDB:GN00141  
 C|Genetics:  
 ;Map position: 1

Query Match	18.4%;	Score 645;	DB 2;	Length 332;
Best Local Similarity	46.4%;	Pred. No. 5e-38;		
Matches 141;	Conservative 45;	Mismatches 80;	Indels 38;	Gaps 6

[illegible]

## RESULT 8

protein F20H12 [imported] - Arabidopsis thaliana  
C|Species: Arabidopsis thaliana (mouse-ear cress)  
C|Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C|Accession: H96627  
R|Theoclygus, A.; Ecker, U.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alons  
Chen, C.W.; Hughes, M.K.; Conn, D.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K  
ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A|Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.; Li, U.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maltli, R.; Marziani  
Cizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.R.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A66141; MUID:21016719; PMID:11130712  
 A:Accession: H96827  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-308 <STO>  
 A:Cross-references: GB:AF005173; NID:97715603; PIDD:AAF68121.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F20B17.12  
 A:Map position: 1

Query Match 17.6%; Score 616; DB 2; Length 308;  
 Best Local Similarity 50.0%; Pred. No. 5e-36;  
 Matches 132; Conservative 36; Mismatches 58; Indels 38; Gaps 7;

QY 254 SSCVTAAPSGTDSVAVDAAKRGHAKLGGQOPVHRKSIDTFPGORTSQVGRVTRHRTGRYA 313  
 DB 18 SOKTASASASIALTSKRRK-----KSPPRNAPL-----QRSSPIRGVT--RRTGRTYA 65  
 QY 314 HLMDSCKEQTGRKGRQVYLGVDMEKAAAYDLAALKYWGSPSTHINFSIENYQVLE 373  
 DB 66 HLMDSKSWNDTQTKKGRQ---GAYEEBAARAAYDLAALKYWGSDTLNFPSPYDEVDK 122  
 QY 374 EMKNSRQRYVAHLRKSQSGFSGASIRGYTRHHOGRMQRARIGVAGNKDLYLGTFT 433  
 DB 123 EMEGSKERYISLRKSSGFSGRYSKRGVAHHNRMARIGRFGNKLYLGYAT 182  
 QY 434 QEEAAEAYVAAIKRGANAATNPDISRYVERIYASSNLAGEIARRKKDNDPR----- 488  
 DB 183 QEEAAIAYDIAIEYRGNAATNPDISRY-----LNPAAADKADSDSKPRSP 231  
 QY 489 ---NKDIDYKSVTVSNNEETVQ 509  
 DB 232 SREPSSDNK---SPKSEEVIE 251

RESULT 9  
 S54116  
 hypothetical protein - maize  
 C:Species: Zea mays (maize)  
 C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Sep-1997  
 C:Accession: S54116  
 R:Daniel, T.J.; Edwards, R.  
 submitted to the EMBL Data Library, January 1995  
 A:Description: Complementation of a heat shock sensitive mutant of *Escherichia coli* defect

A:Reference number: S54116  
 A:Accession: S54116  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-485 <DAN>  
 A:Cross-references: EMBL:247554; NID:91103747; PID:91360703

Query Match 15.2%; Score 534; DB 2; Length 485;  
 Best Local Similarity 33.5%; Pred. No. 6.2e-30;  
 Matches 150; Conservative 48; Mismatches 146; Indels 104; Gaps 13;

QY 187 DCSLMPQMTBGLKXNVAFTRFSTHQVLEBOQNCG---MGNERNVGS-----LG 234  
 DB 4 DMSAYVPH-----HWLSTSLNNYHHGLBFSNSGSPGLGDCGAVESPRVDFDLG 57  
 QY 235 SVGC-GELOSLSMSPGSSCVTAAPSGTDSVAVDAAKRGHAKLGGQ---QVTR-- 286  
 DB 58 GVCVAPAPRSRLQIR-----ITSLAASGCSITAFILHHYPAASGTTGSEPLSRFT 110  
 QY 287 ---KSIDTFGQRTSGYRGVTR-----HRMT-----GRYEAHLMDNSCKEQTGRK 328  
 DB 111 LAAMSTDVAAVESDASGPPRRSASAHSTVYASGTGGRDMRTGNNSCREGSRK 170  
 QY 329 GRQVYLGVDMEKAAAYDLAALKYWGSPSTHINFSIENYQVLEEMKNSRQRYVAHLR 388  
 DB 171 GRQVYLGVDMEKAAAYDLAALKYWGSPSTHINFSIENYQVLEEMKNSRQRYVAHLR 230

QY 389 RKSSGFSRGSATYGVTRHHOGRMQRARIGVAGNKDLYLGTFTSGEAAEAYVAIKF 448  
 DB 231 RKSSGFSRGSATYGVTRHHOGRMQRARIGVAGNKDLYLGTFTSGEAAEAYVAIKF 230  
 QY 449 RGAANAATNPDISRYVERIM-----ASSNLAGEIARR 481  
 DB 291 PRACRHHOLDMSRYDVESILSSDLPVGGASGRAAKFPDLSIQPSGAAAMLAGAAAS 350  
 QY 482 KKNDPRKCDI-----DYNKSVTVSNNEETVQVQAGNNNNEND 520  
 DB 351 QATMPFSKDYMSLLALHYQQQCEQERQFPASAYEAAGSGVAVDFPMGTSSSGNNNTGS 410  
 QY 521 S-EW---KMYLFNHPSSQQQANNGSD 543  
 DB 411 GVMGATSGAVVGQODSSSKQNGYASN 438

RESULT 10  
 T47591  
 aintegument-like protein - Arabidopsis thaliana  
 N:Alternate names: protein T12E18.10  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T47591  
 R:Blocker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Queciet, F.; Salanoubat M. Mewes  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24469  
 A:Accession: T47591  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-205 <BLO>  
 A:Cross-references: EMBL:AL132971  
 A:Experimental source: cultivar Columbia; BAC clone T12E18  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 71/2; 99/1, 128/3; 170/2  
 A>Note: T12E18.10

Query Match 14.8%; Score 517; DB 2; Length 205;  
 Best Local Similarity 55.1%; Pred. No. 2.7e-29;  
 Matches 109; Conservative 21; Mismatches 58; Indels 10; Gaps 4;

QY 241 LQSLSLMSPGSQ--SCVTAAPSGTDSVAVDAAKRGHAKLGGQOPVHRKSIDTFGCR 295  
 DB 5 LTTSTCSSSPSSSVSSSTTSSPIQSEAPPRKRAKSSPSGDKS--HNPTSPASTRK 62  
 QY 236 TSQYGVTRHRTGRYEHLMDNSCKEQTGRKGRQVYLGVDMEKAAAYDLAALKY 355  
 DB 63 SSIYGVTRHRTGRFEHLMDKSSMNSIONKKGKQ---GAYSEEAHAATYDLAALKY 119  
 QY 356 GSTHINFSIENYQVLEEMKNSRQRYVAHLRKSQSGFSGASIRGYTRHHOGRMQR 415  
 DB 120 GPDILNPAATYKLEBQRYKYEYLASLRQSSGFSRYSKRGVAHHNRMARIGRFG 179  
 QY 416 RIGRYAGNKDLYLGTFT 433  
 DB 180 RIGRYAGNKLYLGTFT 197

RESULT 11  
 A85436  
 APTAL2 protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: A85436  
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr  
 Nature 402, 769-777, 1999  
 A:Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A85001; MUID:20083488; PMID:10617198  
 A:Accession: A85436  
 A:Status: preliminary  
 A:Molecule type: DNA



QY 431 FSTCEAAEAYDVAIKFRGANAATNFDISRVDYERIMASSNLLAGELARRKNDPRNK 490  
 DB 275 FDSEVEAARADKALINNGREAVTNFEMSSYQNE-INSESN-----NS 317  
 QY 491 DIDYKSVVTSVNN-----EETVOVQAGNNNNENDSEMKVTLFNHP----SQQ 534  
 DB 318 EIDLNLGISTGNAPKONGRLFHPNSNTYERORG-VSLRIDNEYMKRPVTPPLPYGSSD 376  
 QY 535 QOANGSDOKIMNGYRNSAFSMALQ 562  
 DB 377 HRLYNNGA-----CPSYNNPAGGRATE 398

## RESULT 14

T01574  
 A:Title: APTAL2-like protein Glosy15 - maize  
 C:Species: Zea mays (maize)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
 C:Accession: T01574  
 R:Moore, S.P.; Sisec, P.H.  
 Genes Dev. 10, 3018-3027, 1996  
 A:Title: Glosy15, an APTAL2-like gene from maize that regulates leaf epidermal cell  
 A:Reference number: Z15175; MUID:97115883; PMID:8957002  
 A:Accession: T03981  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-446 <MO>  
 A:Cross-references: EMBL:U41466; NID:G1732030; PID:AA049567.1; PID:G1732031  
 A:Experimental source: strain indred W64A; leaf  
 C:Genetics:  
 A:Gene: Glosy15

Query Match 13.8%; Score 483; DB 2; Length 446;  
 Best Local Similarity 31.9%; Pred. No. 2.2e-26;  
 Matches 119; Conservative 74; Mismatches 114; Indels 66; Gaps 10;

QY 250 PGSQSCVT---APSGTDSVAVDAKKGHAKLQ-----KQVHRKSIDTFGQRTSQY 299  
 DB 55 PDRRAAVTQOPFPPTTAQATMEQCHVAGSABQWVSSASRSRRGRSSQY 114  
 QY 300 RGVTRHRTGRYEALHMDNSCKKEGQTRKRGQVYLGVDMEKAAAPADLAALKTWGST 359  
 DB 115 RGVTFRRTRGRWESHIMD--C-----GKQVYLGFPDTAQAARADVAQAIKFRGANA 164  
 QY 360 HINFESENQVOLEEKNNROEYVAHLRRKSSGFSRGASITRGVTRHHQHRQARIGR 419  
 DB 165 DINFLLDDYDEMKKKDKSKKEFVYVLRQAGFVRGSSRRGVTO--HKCGKMAKRGQ 223  
 QY 420 VAGNKDLYLGTFSTOEAAEAYDVAIKFRGANAATNFDISRVDYERIMASSNLLAGELA 479  
 DB 224 LMGKTVYVGLVYTEREAQAVDKAIKCYKEAVTNFDAQSYD----- 267  
 QY 480 RRKKNDDPRNKDIDYKSVVTSVNNETVQVQA---GNNNNENDSEMKVTLFNHPSSQOQ 536  
 DB 268 -KELQGFWDGDELDELSLGCASSDPTVAVEAFSPATSSSRKQRTMTLTLGLPEEBST 326  
 QY 537 ANNGSDOKIMNGYRNSAFSMALQDILGIDSVSGGNNMDESKIGTHFSNTSLVT 596  
 DB 327 GAG-----YFHPAAGMFGRRPADGHVAVAPPRHQQOQOQ--GQH----- 364  
 QY 597 SLSSREASPEKR 609  
 DB 365 ---AADPAAPER 374

## RESULT 15

T01574  
 A:Title: indeterminate spikelet 1 - maize  
 C:Species: Zea mays (maize)  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 29-Oct-1999  
 C:Accession: T01574  
 R:Chuck, G.; Meeley, R.B.; Hake, S.  
 Genes Dev. 12, 1145-1154, 1998

A:Title: The control of maize spikelet meristem fate by the APTAL2-like gene indeter  
 A:Reference number: Z14353; MUID:98221107; PMID:9553044  
 A:Accession: T01574  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <CHU>  
 A:Cross-references: EMBL:AF048900; NID:G2944039; PID:AA005206.1; PID:G2944040  
 A:Experimental source: strain B73  
 C:Genetics:  
 A:Gene: idsl  
 A:Map position: 1

Query Match 13.3%; Score 467.5; DB 2; Length 433;  
 Best Local Similarity 30.0%; Pred. No. 2.6e-25;  
 Matches 139; Conservative 68; Mismatches 164; Indels 93; Gaps 12;

QY 201 NMVAPTRBPSHQQVLEQOMNGMNERNGVSLGVCGLQSLSLMSPGSQSCVT-- 258  
 DB 6 NVASPADSGTSSSVL-----NSADGGFRFGL-LGSP--VDDDCSGEMAFGASTGFMTRQ 58  
 QY 259 -----APSGTDSVAVDAKKGHAKLQCKQVHRKSIDTFGQR 295  
 DB 59 LFPSPPEPPEPEPVAAPVPWQPPRADLGVQKPVAPAKVTRGP-----RSR 109  
 QY 296 TSQYRGVTRHRTGRYEALHMDNSCKKEGQTRKRGQVYLGVDMEKAAAPADLAALKTW 355  
 DB 110 SSQYRGVTFRRTRGRWESHIMD--C-----GKQVYLGFPDTAQAARADVAQAIKFR 159  
 QY 356 GPSTHINFESENQVOLEEKNNROEYVAHLRRKSSGFSRGASITRGVTRHHQHRQAR 415  
 DB 160 GLDADINFLSDYEDDLKQRMNWKKEFVHILRQSTGFRGSSKRYGVL--HKCGRWEA 218  
 QY 416 RIGRVAGNKDLYLGTFSTOEAAEAYDVAIKFRGANAATNFDISRVDYERIMASSNLLA 475  
 DB 219 RMGOLGAKTIVGLDSEVEAARADRAALRNGREAVTNPEPSSYN-----AGDNNLR 273  
 QY 476 GELARRKKNDDPRNKDIDYKSVVTSVNNETVQVQAGNNNNENDSEMKVTLFNHPSSQOQ 535  
 DB 274 DTEEAIDDDAIDLRLISQPNVODPKRDNTL---AGLQPTCSPSSSNTVASSCPWSSS 330  
 QY 536 QANGSDOKIMNGYRNSAFSMALQDILGIDSVSGGNNMDE----- 580  
 DB 331 SPMP-----GTHQNPVVSFHHORLYSSACHGFPFPHQVQEPVERRPBLGQPP 380  
 QY 581 -----SSKIGTHFSNTSLVTSLSSSREASPEKRGSLLFP 616  
 DB 381 SWMQAGSPHVLHHSAASSGFSIAAGANGKVPFSPHPPACP 424

Search completed: March 9, 2004, 10:47:35  
 Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:31:03 ; Search time 17 Seconds  
(without alignments)  
2030.736 Million cell updates/sec

Title: US-10-024-632-2

Sequence: 1 MKRINESNNTDGNHNLG.....RPSAISLSLHPFASWTD 663

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	517	14.8	432	1	AP2_ARATH
2	156.5	4.5	248	1	PT12_LYCES
3	155	4.4	5560	1	SPEN_DROME
4	147	4.2	248	1	AP23_ARATH
5	141.5	4.0	344	1	RAVI_ARATH
6	139.5	4.0	222	1	ERF4_ARATH
7	134	3.8	243	1	ERF2_ARATH
8	129.5	3.7	724	1	CC11_MOUSE
9	128.5	3.7	300	1	ERF5_ARATH
10	127.5	3.6	1309	1	RAV3_ARATH
11	124	3.5	2482	1	ATRX_PANTR
12	122	3.5	1584	1	KYK1_DICDI
13	122	3.5	725	1	CC11_PANTR
14	121	3.5	2843	1	APC_HUMAN
15	120	3.4	726	1	CC11_HUMAN
16	119.5	3.4	161	1	PT15_LYCES
17	119.5	3.4	268	1	ERF1_ARATH
18	118.5	3.4	918	1	YMB_CABEL
19	118.5	3.4	3075	1	LM1_HUMAN
20	118.5	3.4	1046	1	KYK4_YEAST
21	118	3.4	1444	1	NCOS_HUMAN
22	115.5	3.3	225	1	ERF3_ARATH
23	115	3.3	2492	1	ATRX_HUMAN
24	114.5	3.3	634	1	HNFA_CHICK
25	114.5	3.3	5147	1	PCLO_HUMAN
26	113.5	3.2	984	1	HYSA_STRAS
27	113.5	3.2	730	1	GLN3_YEAST
28	113	3.2	1050	1	LKIA_ACTAC
29	113	3.2	1085	1	CUT7_SCHPO
30	113	3.2	1569	1	GLI3_XENTLA
31	113	3.2	2492	1	ATRX_PONPY
32	113	3.2	592	1	HN1B_XENTLA
33	112.5	3.2			

## ALIGNMENTS

RESULT 1	AP2_ARATH	STANDARD;	PRT;	432 AA.
AC	P47927			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Floral homeotic protein APEAL2.			
GN	AP2 OR ATG36920 OR AP22.49 OR C7A10.440.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
CC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
CC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTANTS AP2-1 AND AP2-5.			
RC	SPRAIN-cv. Columbia; TISSUE=flower;			
RX	MEDLINE=95003709; PubMed=7919989;			
RA	Yotoku K.D., den Boer B.G.W., van Montagu M., Okumura J.K.;			
RT	"Control of Arabidopsis flower and seed development by the homeotic			
RL	gene APEAL2.";			
RM	Plant Cell 6:1211-1225(1994).			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-cv. Columbia;			
RX	MEDLINE=20083486; PubMed=10617198;			
RA	Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,			
RA	Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,			
RA	Harris B., Ansoy W., Brand P., Grivell L.A., Rieger M.,			
RA	Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,			
RA	Kreis M., Delseny M., Puidomech P., Watson M., Schmidheini T.,			
RA	Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,			
RA	Vos P., Hohenstein U., Zimmermann W., Weller H., Ridley P.,			
RA	Langham S.-A., McCallagh B., Bilham L., Robben U., Vandebussche F.,			
RA	Van der Schueren J., Gymnopoulos B., Chang Y.-J., Vandebussche F.,			
RA	Braeken M., Welfens J., Voet M., Bastiaens I., Aert R., Defoor E.,			
RA	Weltzienegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,			
RA	Holzer E., Brandt A., Peters S., van Staveren M., Dirksen W.,			
RA	Moolman P., Klein lanchorst R., Rose M., Haut J., Koester P.,			
RA	Berneser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,			
RA	De Keyser A., Bysmaert C., Gielen J., Villarroel R., De Clercq R.,			
RA	Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray Allen S.,			
RA	Clark L., Doggett J., Hall S., Kay M., Leonard N., McIlroy K., Naves R.,			
RA	Peterson A., Rajendram M.A., Lyne M., Benes V., Rechmann S.,			
RA	Borkova D., Bloeker H., Schaefer M., Grimm M., Lehnert T.-H.,			
RA	Dose S., de Haan M., Maarse A.C., Schaefer M., Mellert-Auer S.,			
RA	Gabel C., Fuchs M., Fattmann B., Grandjean K., Dauner D., Herzl A.,			
RA	Neumann O., Argitlou A., Vitale D., Lignori R., Piravandi E.,			
RA	Wassenaar O., Argitlou A., Vitale D., Lignori R., Piravandi E.,			
RA	Schnabl S., Hiller R., Schmidt W., Leclercq A., Aubourg S.,			
RA	Chedor F., Cooke R., Berger C., Monfort A., Casacuberta E.,			
RA	Gibbons T., Weber N., Vandenbol M., Bagnues M., Terol J., Torres A.,			
RA	Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,			
RA	Heijnen L., Schwarz S., Scholler P., Heber S., Franck P., Bielke C.,			
RA	Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,			
RA	Zaccaria P., Bevan M., Wilson R.K., de la Bastide V., Habermann K.,			

Q24523 drosophila  
Q09591 caenorhabd  
P23499 colutrix co  
Q84498 homo sapien  
P35662 bos taurus  
P53550 saccharomyc  
Q99298 mus musculu  
Q82937 yersinia pe  
Q04233 saccharomyc  
P47136 saccharomyc  
P21613 loligo peal  
Q9hwb8 pseudomonas

RA Parnell L., Dethia N., Gnoj L., Schütz K., Huang E., Spielgei L.,  
RA Sekhon M., Murray J., Shest P., Cordes M., Abu-Threideh J.,  
RA Lonerking T., Kalcik J., Graves T., Hamon G., Edwards J.,  
RA Lareelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Mux P., Bentley D., Fulton L., Miller N., Greco T., Kemp K.,  
RA Kramer J., Platten L., Mardis E., Dante W., Pepin K., Hillier L.W.,  
RA Nelson J., Sleeth J., Ryan E., Andrews S., Geisel C., Layman D.,  
RA Du H., Ali T., Bergoff A., Jones K., Drone K., Cotton M., Joshi C.,  
RA Antoniou B., Zidanik M., Strong C., Sun H., Lamar B., Jordan C.,  
RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman T., Tili S.,  
RA Grady S., Shohy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,  
RA Chen E., Marra M.A., Marliensen K., McCormie W.R.,  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RT thaliana." ;  
RT Nature 402:769-777(1999).  
RL [3]  
RN FUNCTION.  
RP MEDLINE=91256319; PubMed=1675158;  
RX Drews G.N., Bowman J.L., Meyerowitz E.M.;  
RA "Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the  
RT AP2/AP2A2 product." ;  
RT Cell 65:991-1002(1991).  
RL -1- FUNCTION: Probable transcriptional activator that promotes early  
CC the floral meristem identity. Is required subsequently for the  
CC transition of an inflorescence meristem into a floral meristem.  
CC plays a central role in the specification of floral identity,  
CC particularly for the normal development of sepals and petals in the  
CC wild-type flower. Acts as a class cadasteral protein by repressing  
CC the C class floral homeotic gene AGAMOUS in association with  
CC others repressors like LEM1G and SEUSS. It is also required  
CC during seed development.  
CC -1- SUBUNIT: May form homodimer.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- TISSUE SPECIFICITY: Sepals, petals, stamens, carpels, developing  
CC ovules. Inflorescence stem, leaf, and stem.  
CC -1- DEVELOPMENTAL STAGE: It is detectable at low levels throughout the  
CC shoot apex, and at enhanced levels in the inflorescence meristem,  
CC young floral buds, and throughout the early stages of flower  
CC development and organogenesis. During floral organ differentiation  
CC it becomes spatially restricted to specific organ, tissue, and  
CC cell types within the flower.  
CC -1- INDUCTION: Negatively regulated by the C class floral  
CC homeotic protein AGAMOUS in stamens and carpels.  
CC -1- MISCELLANEOUS: Mutations in the AP2/AP2A2 gene result in the  
CC ectopic expression of AGAMOUS, leading to the replacement of  
CC sepals by carpels and stamens and of petals by stamens.  
CC -1- SIMILARITY: Contains 2 AP2/ERF domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation in  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce>  
CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
-----  
DR EMBL: U12546; AAC3770.1; -  
DR EMBL: Z99707; CAB16765.1; -  
DR EMBL: AL161590; CAB80358.1; -  
DR PIR: A85436; A85436.  
DR TRANSFAC: T017741;  
DR InterPro: IPR001471; TF\_ERF.  
DR Pfam: PF00847; AP2-domain; 2.  
DR PRINTS: PR00367; ETRHSPELEMT.  
DR ProDom: PD001423; TF\_ERF\_1.  
DR SMART: SMC0380; AP2; 2.  
KM Flowering: Transcription regulation; Activator; Developmental protein;  
KM Nuclear protein; DNA-binding; Repeat.  
FT DOMAIN 14 50  
FT 119 128  
FT 128 192  
FT DNA\_BIND 128 192  
FT AP2/ERF 1.  
DOMAIN 119 128  
DNA\_BIND 128 192  
AP2/ERF 1.  
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
INVOLVEMENT WITH TRANSCRIPTION).  
ASP/GLU/SER-RICH (ACIDIC; POTENTIAL).  
NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

Query Match	14.8%	Score 517	DB 1	Length 432
Best Local Similarity	31.3%	Pred. No. 1,4e-28		
Matches 163	Conservative	66	Mismatches 150	Indels 142
				Gaps 19
FT DNA_BIND	220	286	AP-2/ERF 2	
FT MYOAGIN	159	159	G-2E: TEMPERATURE SENSITIVE; AP2-5.	
FT MYOAGIN	251	251	G-2S: TEMPERATURE SENSITIVE; AP2-1.	
FT MYOAGIN	420	420	O-2E: TEMPERATURE SENSITIVE; AP2-5.	
SO SEQUENCE	432 AA	47833 MW	376A355291EFPB42 CRC64	
QY	171	YQAFLEBETKETH-----VSDSSLMPTQTEGLK---NWVAPTEPSTHQVLE	217	
Db	9	HQGRRESESEFVCSSPKRVGSPFSSSSSAVVIEDSDDELNRVFPNNPLVTHQFPE	68	
QY	218	QQNNQCNGBNRNGVSLSSVGGGLQSLSLKMSPS-----OSSCVTAPSGTDSVAYD	269	
Db	69	MDNSNG-----GVASG-----FPFAHWFGKFCOSDLATGSSAGKATNVA	108	
QY	270	AKRGAHAKGQKQFVHRKSIDTFQORTSOYRGVTRHMTGRYEALWDNSCKEGQTRKG	329	
Db	109	A-----AVBEAQPL-KKSRGRPFRSSQYRGVTFYRTGRMBEHIWD--C-----G	152	
QY	330	ROYVLGSDYDEKKAARYDALAKYMPSTHINSIENYQVQLEEMKMSROEYVAHLR	389	
Db	153	KQYVLGGDTTAAARAAYDRALAFRGVEADININIDYDDDLQMNNTLDEEVYHLR	212	
QY	390	KSGSFSRSGASIVRGVTRHHQGRNQASIGRVAGNKDLYLGTSTQEEAAEYDVAATKFR	449	
Db	213	QSTGPPOSSKYRGVTL-HKCGWEANMGFLGKYYVYLGLFDEVEAARAAYDKAAIKCN	271	
QY	450	GAAVTPNDIRYQVERIMASSNLLAGELARKDNDPRNNDIDYNSVYTVSANNBETVQ	509	
Db	272	GKDAYTFDPSTIYDEBLNASSG-----NPTTPQDNLN-----	305	
QY	510	VQAGNNNNENDSEWKMYLFNHPSSQOQANGSGDQIKMNCGRYNSAFSMALODLIGIDS	569	
Db	306	LSLGNANS-----KKKSQDMRLRMNQOOD-----SLHSNEVLGIGQ	343	
QY	570	VGSGQH-----NMLDESKIGT--HFS-----NTSSLVTSLSSSREASPEK	608	
Db	344	TGMLNHTNNSHQPPGSSNISGGGFLSFPAAENHRFPGRASTQVLTNMAASSGFSFHH	403	
QY	609	RGPSLLFPMPMETKIYNPIGTSVTSMUPSPVQ--NRPS 646		
Db	404	-----HNQIFNSTSTPHQNMVLQTFQFPILMRPS 432		
RESULT 2				
PT16_LYCES	STANDARD	PRT	248 AA	
AC	004682			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Pathogenesis-related genes transcriptional activator PT16.			
OS	Lycopersicon esculentum (Tomato).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			
OC	lamiales; Solanales; Solanaceae; Solanum.			
OX	NCBI_TaxID=4081;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97357308; PubMed=9214637;			
RA	Zhou J., Tang X., Martin G.B.;			
RT	"The Pto kinase conferring resistance to tomato bacterial speck			
RT	disease interacts with proteins that bind a cis-element of			
RT	pathogenesis-related genes.";			
RL	EMBO J. 16:3207-3218(1997).			
CC	-1- FUNCTION: Transcription factor that binds to the GCC-box			
CC	pathogenesis-related promoter element. Activates plants defense			
CC	genes			
CC	-1- SUBUNIT: Interacts with the Pto kinase.			

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL: U89257; AAC49741.1; -  
 DR PIR: T07728; T07728.  
 DR HSSP: O80337; 2GSC.  
 DR TRASPAC: T04717; -  
 DR InterPro: IPR01471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 1.  
 DR PRINTS: PR00367; ETRHSPLENT.  
 DR ProDom: PD001423; TF\_ERF; 1.  
 DR SMART: SM00380; AP2; 1.  
 DR Transcription regulation: DNA-binding; Activator; Nuclear protein.  
 KW DNA\_BIND 95 159 AP2/ERF  
 FT SEQUENCE 248 AA; 27913 MW; 115BD44DEB1B992 CRC64;  
 SQ

Query Match 4.5%; Score 156.5; DB 1; Length 248;  
 Best Local Similarity 29.0%; Pred. No. 0.00072;  
 Matches 47; Conservative 16; Mismatches 54; Indels 45; Gaps 4;

QY 359 THINSIENYQVQLEMKMSKQVEYVAALRRKSSFGASITGVTRHHGHWQRIG 418  
 DB 67 TEINIMPSTKSGIDRRKRSVSDVT--RRKK-----FQGV-RQPPGMALEIR 114  
 QY 419 RVAQNKDYLGTFTSTOEAAEYVAALFRGANAVTFDSRYVERIMASNLAGEI 478  
 DB 115 DTRGRKRWLGTYPPEBAAYVDKAAVKLQPDVNTFPVST----- 157

QY 479 ARRKQNDPRNDIDYNSVTVSNNEETVQVQAGNNNEED 520  
 DB 158 -----TAEVTVTETETESVADGDKSEND 183

RESULT 3  
 SPEN\_DROME STANDARD; PRT; 5560 AA.  
 AC Q8SX83; Q9NHN1; Q9NVJ7; Q9UCG3; Q9VPL1; Q9VPL2;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Split ends protein.  
 GN SPEN OR CG18497.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4), FUNCTION, ALTERNATIVE PROMOTER  
 RP USAGE, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND MUTANTS E9 AND  
 RP D57.  
 RC TISSUE=Embryo;  
 RX MEDLINE=20025936; PubMed=10556062;  
 RA Welleite E.L., Harding K.W., Mace K.A., Ronehaugen M.R., Wang F.Y.,  
 RA McGinnis W.;  
 RT 'Spn encodes an RNP motif protein that interacts with Hox pathways  
 RT to repress the development of head-like sclerites in the Drosophila  
 RT trunk.';  
 RL Development 126:5373-5385(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 4).  
 RC TISSUE=Embryo;  
 RX MEDLINE=2017049; PubMed=10655223;  
 RA Rebay T., Chen F., Hsiao F., Kolodziej P.A., Kuang B.H., Laverly T.,  
 RA Sun C., Voas M., Williams A., Rubin G.M.;

RT "A genetic screen for novel components of the Ras/mitogen-activated  
 RT protein kinase signaling pathway that interact with the Yan gene of  
 RT Drosophila identifies split ends, a new RNA recognition motif-  
 RT containing protein.";  
 RL Genetics 154:695-712(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RC TISSUE=Embryo;  
 RX MEDLINE=20171275; PubMed=10704397;  
 RA Kuang B.H., Wu S.C.-Y., Shin Y.-A., Luo L., Kolodziej P.A.;  
 RT "Split ends encodes large nuclear proteins that regulate neuronal  
 RT cell fate and axon extension in the Drosophila embryo.";  
 RL Development 127:1517-1529(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams W.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abri'l J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Bencs P.V., Bernan B.P., Bhattacharya P., Brothier P.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockton P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Foster A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Hewland T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
 RA Lascko P., Lei Y., Levitsky A.A., Li U.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Malshina N.V., Mobarry C., Morris J., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson X.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach U.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [5]  
 RP REVISIONS, AND ALTERNATIVE SPLICING.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bencicourt B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.F.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [6]  
 RP SEQUENCE OF 424-2002 FROM N.A.  
 RC STRAIN=Berkeley; TISSUE=Embryo;  
 RX MEDLINE=22426066; PubMed=12537569;



RA Stapleton R.A., Carlson J.W., Brokstein P., Yu C., Champe M.,  
RA George R.A., Guarini H., Krommiller B., Paclob J.M., Park S., Wan K.H.,  
RA Rubin G.M., Ceiniker S.E.,  
RA "A Drosophila full-length cDNA resource."  
RT Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
RN [7]  
RP FUNCTION.  
RX MEDLINE=20253107; PubMed=10790398;  
RA Lamer M.E., Elend M., Heidmann D., Herr A., Marzock S., Herzig A.,  
RA Lehner C.F.,  
RT "A screen for modifiers of cyclin E function in Drosophila  
RT melanogaster identifies Cdk2 mutations, revealing the insignificance  
RT of putative phosphorylation sites in Cdk2."  
RL Genetics 155:233-244(2000).  
RN [8]  
RP FUNCTION ON EGF RECEPTOR PATHWAY.  
RX MEDLINE=20414403; PubMed=10959845;  
RA Chen F., Rebay I.,  
RT "Split ends, a new component of the Drosophila EGF receptor pathway,  
RT regulates development of midline glial cells."  
RL Curr. Biol. 10:943-946(2000).  
RN [9]  
RP FUNCTION ON Wg PATHWAY.  
RX MEDLINE=22668876; PubMed=12783785;  
RA Lin H.V., Doroquez D.B., Cho S., Chen F., Rebay I., Cadigan K.M.,  
RT "Split ends is a tissue/promoter specific regulator of WINGLESS  
RT signaling."  
RL Development 130:3125-3135(2003).  
CC -1- FUNCTION: Probable corepressor protein, which regulates different  
CC key pathways such as the EGF receptor and Wg pathways. Involved in  
CC neuronal cell fate, survival and axon guidance, cell cycle  
CC regulation and repression of head identity in the embryonic trunk.  
CC May act with the Hox gene Deformed and the EGF receptor signaling  
CC pathway. Positive regulator of the Wg pathway in larval tissues  
CC but not in embryonic tissues. May act as a transcriptional  
CC corepressor protein, which repress transcription via the  
CC recruitment of large complexes containing histone deacetylase  
CC proteins.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Comment=2 isoforms, 1 (shown here) and 2, are produced by use of  
CC alternative promoters;  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=1;  
CC IsoId=Q8SX83-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q8SX83-2; Sequence=VSP\_008565; VSP\_008566;  
CC Note=No experimental confirmation available;  
CC Name=3; Synonyms=Spenu;  
CC IsoId=Q8SX83-3; Sequence=VSP\_008567;  
CC Note=Produced by alternative splicing of isoform 1;  
CC Name=4; Synonyms=Spens;  
CC IsoId=Q8SX83-4; Sequence=VSP\_008565; VSP\_008566; VSP\_008567;  
CC Note=Produced by alternative splicing of isoform 2;  
CC -1- TISSUE SPECIFICITY: Ubiquitous. Expressed prior to cellularization  
CC in stage 3 embryos, and in blastoderm cells, including pole cells.  
CC Expressed throughout the rest of embryogenesis. Later, it is  
CC expressed at higher level in epidermal cells and CNS.  
CC -1- DEVELOPMENTAL STAGE: Isoform 3 is expressed both maternally and  
CC zygotically.  
CC -1- SIMILARITY: Belongs to the Spen family.  
CC -1- SIMILARITY: Contains 3 RNA recognition motif (RRM) domains.  
CC -1- SIMILARITY: Contains 1 SPOC domain.  
CC -1- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; AF188205; AAF13318.1; -  
DR EMBL; AF184612; AAF26299.1; -  
DR EMBL; AF221715; AAF34661.1; ALT\_INIT.  
DR EMBL; AE003590; AAP51334.2; -  
DR EMBL; AE003590; AAP51335.2; -  
DR EMBL; AE003590; AAF10511.1; -  
DR EMBL; AY094789; AAF11411.1; ALT\_SEQ.  
DR HSSP; P09651; 1HAI.  
DR FlyBase; FBgn0016977; spen.  
DR GO; GO:0007411; P:axon guidance; IMP.  
DR GO; GO:0008347; P:glia cell migration; IMP.  
DR InterPro; IPR00504; RNA\_rec\_moc.  
DR Pfam; PF00076; Rrm; 3.  
DR SMART; SM00360; RRM; 3.  
DR PROSITE; PS50102; RRM; 3.  
DR PROSITE; PS00030; RRM\_RNP\_1; FALSE\_NEG.  
DR PROSITE; PS50917; SPOC\_1;  
KW Transcription regulation; Repressor; Developmental protein;  
KW Nuclear protein; Repeat; RNA-binding; Coiled coil;  
KW Alternative promoter usage; Alternative splicing.  
FT DOMAIN 554 632 RNA-BINDING (RRM) 1.  
FT 554 730 RNA-BINDING (RRM) 2.  
FT DOMAIN 656 730 RNA-BINDING (RRM) 3.  
FT DOMAIN 734 806 RNA-BINDING (RRM) 3.  
Query Match 4.4%; Score 155; DB 1; Length 5560;  
Best Local Similarity 19.7%; Pred. No. 0.081;  
Matches 144; Conservative 85; Mismatches 266; Indels 236; Gaps 30;  
CC -----  
QY 7 SNTDGGNNNNWLGSLSPMKMEATSAATVPTFTYMSPSQSHLSPFGCYGNGNTH 66  
DB 358 TNNFSNNSNNTTATPTPTIAGGAG-----SVGLSGAAGAGCSNSG----- 404  
QY 67 SPLTVMLKSDGSLCILEAKRSQTVMTSSPKLEDFGATMGHEGHERGLSD 126  
DB 405 -----TSGDLNAAVLAANDGVTHPIRIRNLHGRT--TSSRSRSRSPSSV 455  
QY 127 SIYVNSQNAEOPNRDLISQPFROGHMSVQTHPYVSGLAGHGLYQAPLEETTKETHVS 186  
DB 456 SSSSSSSSSSHSSSHSSSHSSSPVQSG----- 481  
QY 187 DCSLSQMTGMEGLK---NMYAPTRESTHQVLEQGMNGMNERNGVSLSGVGGELQ 242  
DB 482 NCA-----MAEGSSRTVNSVTYVTSNNSPSGTAIVVSRAGVG-----GGGS-S 525  
QY 243 SLTSLMSPGSSQSCVTA-----PSGTDVAVDARKKHAK 277  
DB 526 SSSSSSSSSSSGSLTANPVVHSEDNRPLAIRVNLPARSSDLSKDGLFHEYKKGKVT 585  
QY 278 LGQKQPVHRKSIDTFGQRTSQYRGVTHRWGTGYEALMNSC-KKGGQTRKGRQV--- 332  
DB 586 W-----VKVVGQNSERYALV-----CFKKPDVEKALEVSHDK 618  
QY 333 -YLG-----GYDNEEYAARAYDLAALKYMGPSSTHINFSIENYOVQLEMKMSROE 382  
DB 619 HFFGCKIEVERGYQVDYEDNEDEFRYELDEHYHKSRTLP-IGNLE-----KDIAGE 671  
QY 383 VVAHLRR-----KSGG-----PSRGASITYGVTGTR---HHQGMQARIGRA 421  
DB 672 LRSHFEAFGEIIEIDIKKGIMAAVAFQYSDIVSVVAMKMGDEHLSNRILKGFQSKM 731  
QY 422 GNKLYLGTFSTOEAEAVYDAIKRFGANAVTNFDIR-----YD-VERIMASSN 472  
DB 732 PTNCWMD--GVDRKVESF--LOSQTRRGATYKVISIDNRQLALVLYQVOQAQAAVK 787  
QY 473 LLAELARKKDNDPRNKDIDYKSVTVSVNNEVTVQAGNNNENDESKVTLFPHPS 532  
DB 788 DMRGTLIRKK-----LQVDPAARECD-----AFYDQ 816  
QY 533 QQQQANGSGDQKIMNCYNSAFSMALODLIGDVGSGQHMLDESSKIGT----- 586  
DB 817 EKQQQSSGSNPRF---SRYSSASS--LQSRASSFSRHHNNNSNDCCPINTPGASS 871

QY 587 HFSNTSLVTSLSSEASPEKRGPSLLPMPMETKIVN-----PIGTSVTSMLP 637  
 DB 872 GISSASLILNG-SISINISINIGNACSNAMPASLAASVSCNNAAGTVPASTSPSGVS 930  
 QY 638 SPTVQMRPSPA 648  
 DB 931 SSSSSLPMSPA 941  
 RESULT 4  
 AP23\_ARATH STANDARD; PRT; 248 AA.  
 AC P47367.023105;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE AP2 domain transcription factor RAP2.3 (Related to AP2 protein 3)  
 GN (Cadmium-induced protein AS30).  
 GN ATG16770 OR MG6.1 OR MG6.24.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Choi S., Baek E., Lee S.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=97338147; PubMed=9192694;  
 RA Okemuro J.K., Caster B., Villarreal R., van Montagu M., Jofuku K.D.;  
 RT "The AP2 domain of APTPA2 defines a large new family of DNA binding  
 RT proteins in Arabidopsis".  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:7076-7081(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA MEDLINE=20277480; PubMed=10819329;  
 RA Sato S., Nakamura Y., Kaneo T., Katoh T., Asamizu E., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
 RT clones".  
 RL DNA Res. 7:131-135(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheek R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.W., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng C.H., Akiyama K., Ansari Y.,  
 RA Aizawa T., Barni J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Ebnji A., Goldsmith A.D., Gurjani M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kaniva A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Saito M., Tamse R., Vaysberg M., Wallender E.K., Kong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis K.W., Teodorick A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome".  
 RL Science 302:842-846(2003).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Trukhan M., Alexandrov N., Lu Y.-F., Flavell R.,  
 RA Feldmann K.A.;  
 RT "Full-length cDNA from Arabidopsis thaliana".  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Nucleus (Probable).  
 CC -1- INDUCTION: By cadmium.  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.

CC -1- CAUTION: Ref.1 sequence is most probably a cloning artifact and  
 CC differs from that shown due to an insertion into the sequence.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 CC EMBL; Z37504; CAA5734.1; ALT\_SEQ.  
 DR EMBL; AF003096; AAC49769.1; -  
 DR EMBL; AB022217; BAB02769.1; -  
 DR EMBL; AY059917; AAL24399.1; -  
 DR EMBL; AY035100; AAK59605.1; -  
 DR EMBL; AY142562; AAN13131.1; -  
 DR EMBL; AY087489; AAM65031.1; -  
 DR HSSP; O60337; 2GCC.  
 DR TRANSFAC; T02658; -  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; ETRSPLEMMNT.  
 DR PRODOM; PD001423; TF\_ERF; 1.  
 DR SMART; SM00380; AP2; 1.  
 KW Transcription regulation; DNA-binding; Nuclear protein; Cadmium.  
 FT DNA BIND 76 141  
 FT DOMAIN 140 143  
 FT POLY-PRO.  
 SQ SEQUENCE 248 AA; 27758 MW; 317D50CC514782F1 CRC64;  
 Query Match 4.2%; Score 147; DB 1; Length 248;  
 Best local similarity 25.1%; Pred. No. 0.0033;  
 Matches 67; Conservative 34; Mismatches 100; Indels 66; Gaps 10;  
 QY 328 KGRQVVLGGYDMEERAAAYDLAAKMG-----PSTHINPSIENYQVLEMKMS 379  
 DB 18 KGRKL-----TAEELMSLELDASADDFWGFYSTSKLHPTNVN-----VKEAIVKKEQ 65  
 QY 380 ROEYVAHLRRKSGSGFSGASIRYGVTRHCHGRMWARIGRAVAKNDLYLGTFSOEBAE 439  
 DB 66 ATEPGRRRRX-----NVYRGI-RKRPWKMAAEIRDRKGRVWLGFNTAEERAM 116  
 QY 440 AYVVAALIKRGAAATNPISRYDYERIMASNLNLAGLARKKDNDRNDIYKSVY 499  
 DB 117 AYVAAKQIRGDKAKLNPFDLHHP-----PPNYTPPPSSPRSTDPARK-----V 163  
 QY 500 TSYNNEETVOVAGNNNNNENDESEWMLFNHPSQQOQANANGSDQKIMNCG-----NYRN 554  
 DB 164 CVV-----SQSESE-----LQSPFVEICIGFONGDEFQNLVGFEPDYDL 204  
 QY 555 SAFSMALQDLIDISVSGQHMLDES 581  
 DB 205 KQOISLSLEFLDLGNTAQPQLDES 231  
 RESULT 5  
 RAVI\_ARATH STANDARD; PRT; 344 AA.  
 AC G92M99;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE DNA-binding protein RAV1.  
 GN RAV1 OR ATG13360 OR Te04.2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seedling;  
 RL MEDLINE=99081843; PubMed=9862967;

RA Kagaya Y., Ohmura K., Hattori T.;  
 RT "RAV1, a novel DNA-binding protein, binds to bipartite recognition  
 RT sequence through two distinct DNA-binding domains uniquely found in  
 RT higher plants";  
 RL Nucleic Acids Res. 27:470-478(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=21016719; PubMed=1130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federici N.A., Kaul S.,  
 RA White O., Alonso J., Altfeld H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chan O., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Egu P., Feldblum I.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Klaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kuritz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Miltcher J., Miranda M., Nguyen M., Nieman W.C., Osborne B.L.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Talon L.J., Tambunga G., Tortum M.J., Town C.D.,  
 RA Utecher T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter U.C., Davis R.W.,  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:816-820(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newman G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Tortum M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carrinci P.,  
 RA Chao Q., Choy N., Egu P., Goldsmith A.D., Guraj M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Kaneko M.,  
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamita A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamme R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome";  
 RL Science 302:842-846(2003)  
 CC -1- FUNCTION: Binds specifically to bipartite recognition sequences  
 CC composed of two unrelated motifs, 5'-CAACA-3' and 5'-CACTG-3'.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined: roots,  
 CC rosette leaves, cauline leaves, inflorescence stems, flowers and  
 CC siliques. Highest expression in roots and rosette leaves. Very low  
 CC expression in flowers.  
 CC -1- DOMAIN: Contains two distinct DNA-binding domains. One is located  
 CC in the N-terminal region and binds to the 5'-CAACA-3' motif. The  
 CC second is located in the C-terminal region and binds to the  
 CC 5'-CACTG-3' motif.  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).  
 CC -----  
 DR EMBL: AB013886; BAA34250.1; -  
 DR EMBL: AC011810; AAC09554.1; -  
 DR EMBL: AY063855; AAL36211.1; -  
 DR EMBL: AY091291; AAM14230.1; -

DR PIR: T51329; T51329.  
 DR HSSP: 080337; 2GCC.  
 DR InterPro: IPR003340; TF\_B3.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain1; 1.  
 DR Pfam: PF02362; B3; 1.  
 DR PRINTS: PR00367; ETRSEPELEMT.  
 DR PRODOM: PD001423; TF\_ERF; 1.  
 KM DNA-binding; Nuclear protein.  
 FT DNA BIND 59 121  
 FT DOMAIN 145 338  
 FT B3.  
 SQ SEQUENCE 344 AA; 38597 MW; 7349B640B305823 CRC64.  
 Query Match 4.0%; Score 141.5; DB 1; Length 344;  
 Best Local Similarity 23.4%; Pred. No. 0.013;  
 Matches 61; Conservative 37; Mismatches 86; Indels 77; Gaps 10;  
 QY 379 SROEYVAHLRRKSSGFS-----RGASIVGTVRHQHGRFCARIGVAGN 423  
 DB 25 AKSSVGNLVRMGSGSVVLDSNGVYAESRKLPSKYGAV-PPNGRMGMOI--YEKH 81  
 QY 424 KDLYLGFSTQGEAAEYDPAKFRGANAYVFDISRDVERIMASSVLLAGEIARRK 483  
 DB 82 QRVWLGFNEDEDAARAYDVAFRRFRDAVYVNFK---DV-----KM 120  
 QY 484 DNDPRNDIDYNSKSVTVSVNNEETVGVQAKNNNNENDESEWKKVLFNHPQQQANGSGD 543  
 DB 121 DED---EVDP---LNSHKSSEIVMLKRTYNELESGK-----RRRNGNM 162  
 QY 544 QKIMCNGNYNSAFSMAUDLDIGDSVSGQHMLDSESKITHTSNTSLYTSISSRE 603  
 DB 163 TRLLTSLGSLND-----GVSTGFRSAALFEKAVTPSDVCKNRLVI----- 205  
 QY 604 ASPEKRGPSLLFMPPEWETKI 624  
 DB 206 ---PKHHAHGHPFLPSNSVS 223  
 RESULT 6  
 ERF4\_ARATH STANDARD; PRT; 222 AA.  
 AC 080340;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ethylene responsive element binding factor 4 (ATERF4).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20181733; PubMed=10715325;  
 RA Fujimoto S.Y., Ohta M., Usui A., Shinishi H., Ohme-Takagi M.;  
 RT "Arabidopsis ethylene responsive element binding factors act as  
 RT transcriptional activators or repressors of GCC box mediated gene  
 RT expression";  
 RL Plant Cell 12:393-404(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=21016720; PubMed=1130713;  
 RA Salanoubat M., Lemke K., Rieger M., Ansoorge M., Unseid M.,  
 RA Fattmann B., Valle G., Bloeker H., Perez-Alonso M., Obermayer B.,  
 RA Delserny M., Boutry M., Grivell L.A., Maché R., Pulsdomech P.,  
 RA De Simone V., Choitane N., Artiguenave F., Robert C., Broclet P.,  
 RA Wincker P., Cactolico L., Weissbach J., Saurin W., Quetier F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Wurmbech E., Drzonek H., Erle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voss H., Holland K., Brandt F., Nyakatura G.,  
 RA Vezzi A., D'Angelo W., Pallavicini A., Toppo S., Simionati B.,

RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,  
 RA Reichelt J., Scharte M., Schoen O., Barges M., Terol J., Clement J.,  
 RA Navarro P., Collado C., Perez-Perez A., Ockenwelder B., Duchem D.,  
 RA Cooke R., Landie M., Berger-Llauró C., Punelle B., Masu D.,  
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cortes A., Casacuberta E.,  
 RA Montfort A., Argirou A., Flores M., Liguori R., Vitale D.,  
 RA Mannhaupt G., Haase D., Schoof H., Rüd S., Zaccaria P., Mewes H.-N.,  
 RA Meyer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo M., Wals A., Utterback T., Fujii C.Y., Shea T.P.,  
 RA Creasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,  
 RA Pat G., Miltescher J., Sellers P., Gill J.E., Feldblum T.V.,  
 RA Pruss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,  
 RA Pruss C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
 RA Sasamoto S., Kimura T., Ideawa K., Kawashima K., Kishida Y.,  
 RA Kiyokawa S., Kohara M., Matsumoto M., Matsuno A., Muraki H.,  
 RA Nakayama S., Nakazaki N., Shimpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.,  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 thaliana.",  
 RT thaliana.",  
 RT Nature 408:820-822(2000).  
 CC -1- FUNCTION: Transcription factor that binds to the GCC-box  
 pathogenesis-related promoter element.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB008106; BAA32421.1; -  
 DR EMBL: AP000413; BAB02150.1; -  
 DR PIR: T52019; T52019.  
 DR HSSP: O80337; 2GCC.  
 DR TRANSFAC: T04636;  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 1.  
 DR PRINTS: PR00367; ETRSPPELEMT.  
 DR PRODOM: PD001423; TF\_ERF; 1.  
 DR SMART: SM00380; AP2; 1.  
 KM Transcription regulation; DNA-binding; Nuclear protein.  
 FT DNA BIND 22 86 AP2/ERF  
 FT DOMAIN 132 137 POLY-GLY  
 SQ SEQUENCE 222 AA; 23741 MW; CEF16ABA5ACFB073 CRC64;  
 Query Match 4.0%; Score 139.5; DB 1; Length 222;  
 Best Local Similarity 27.2%; Pred. No. 0.0093;  
 Matches 67; Conservative 22; Mismatches 82; Indels 75; Gaps 11;  
 QY 401 YRGVTHHGHGQWQARIGRAGKND-LVGTSTQGEAEADVAIFRGANVTN-- 457  
 Db 25 YRGV-RKRPGRPAAL-RDPGRKTRVWLTGDTLAEAPADYADARPRGRKAKTNP 82  
 QY 458 -DISRYVERINASSNLLAGELARRKNDPRNDIDNKSVTVSNNEETVOYQAGNN 515  
 Db 83 FLELSQKVPYTPGARSP-----SQSTLDDCASPTLLVPSATA 120  
 QY 516 NNENDESWKMYVLNHNSSQOQANGNSDQKIMNCGYRNASFMSALQ-----PLIGDS 569  
 Db 121 GNV-----PPQELSLGSGG-----GSCQIWMSPVYFLDMGIGN 158  
 QY 570 VSGGQHNMIDESSK-----IGHFNTSSLVTSLSRRASPEKRPSTLPEM- 617  
 Db 159 VGRGQPPYTSARFSRVVAVATMACGASDSDSSVDF--EGMEKRSQLLIDLIN 214  
 QY 618 --PME 621  
 Db 215 LPPPS 220

RESULT 7  
 ERF2 ARATH STANDARD; PRT; 243 AA.  
 ID ERF2 ARATH  
 AC 080338;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT Ethylene responsive element binding factor 2 (ATERF2).  
 DE ERF2 OR ERF-2 OR AT5G47220 OR MGL5.7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20181733; PubMed=10715325;  
 RA Fujimoto S.Y., Ohira M., Usui A., Shinsai H., Ohme-Takagi M.;  
 RT "Arabidopsis ethylene responsive element binding factors act as  
 RT transcriptional activators or repressors of GCC box mediated gene  
 RT expression.",  
 RL Plant Cell 12:393-404(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Karoh T., Asamizu E., Kotani H.,  
 RA Tsubata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty P1 and PAC  
 RT clones.",  
 RL DNA Res. 7:31-63(2000).  
 CC -1- FUNCTION: Transcription factor that binds to the GCC-box  
 pathogenesis-related promoter element.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB008104; BAA32419.1; -  
 DR EMBL: AB018117; BAA97155.1; -  
 DR PIR: T51989; T51989.  
 DR HSSP: O80337; 2GCC.  
 DR TRANSFAC: T04634;  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 1.  
 DR PRINTS: PR00367; ETRSPPELEMT.  
 DR PRODOM: PD001423; TF\_ERF; 1.  
 DR SMART: SM00380; AP2; 1.  
 KM Transcription regulation; DNA-binding; Nuclear protein.  
 FT DOMAIN 23 26 POLY-GLY  
 FT DNA BIND 114 179 AP2/ERF  
 FT DOMAIN 192 206 POLY-SER.  
 SQ SEQUENCE 243 AA; 26797 MW; C9A4C9791249B5D4 CRC64;  
 Query Match 3.8%; Score 134; DB 1; Length 243;  
 Best Local Similarity 28.6%; Pred. No. 0.026;  
 Matches 60; Conservative 30; Mismatches 82; Indels 38; Gaps 8;  
 QY 302 VTRRWKTRRYEHLWLN-----SCKKEQTRKQRYVIGYDMEKAR--AYDLA--A 351  
 Db 17 ITRHLGGGGENELRLNSTSSCTFS-----WGGFLKENDSEMLVGLKDA 67  
 QY 352 LKYGSPSTHINSIENYVOLEMKMNSROEYVAHLRKSGFERGSIYRGYTRHHQHG 411  
 Db 68 FHFDTSSDLSCLDFPRAVKVETENFAMEEKRAKAIPLYETLVAKHYRGV-RORPWG 126



[illegible]

RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=2018125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asanizu E., Kotani H.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
 RT clones";  
 RL DNA Res. 7:31-63(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Maranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh U., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Guiral M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Ropes C., Hsu H.Y., Iida K., Karnes M.,  
 RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamita A., Meyers C., Ishida M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamse R., Vaysberg M., Walender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome";  
 RL Science 302:842-846(2003).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
 RA Nakajima M., Enju A., Kamita A., Narusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinozaki K.;  
 RT "Arabidopsis thaliana full-length cDNA";  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Transcription factor that binds to the GCC-box  
 CC pathogenesis-related promoter element.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 1 AP2/ERF domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB008107; BAA32422.1; -  
 DR EMBL; AB018117; BAA97157.1; ALT\_INIT.  
 DR EMBL; AF385709; AAK60301.1; -  
 DR EMBL; AY078014; AAL77715.1; -  
 DR EMBL; AK117368; BAC42229.1; -  
 DR F1R; T52020; T52020.  
 DR HSP; O80337; 26CC.  
 DR TRANSFAC; T04637; -  
 DR INTERPRO; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; ETRSPPELENT.  
 DR PRODOM; PD001423; TF\_ERF; 1.  
 DR SMART; SMD00380; AP2; 1.  
 DR Transcription regulation; DNA-binding; Nuclear protein.  
 FT DOMAIN 51  
 FT DNASIS 153 218 AP2/ERF.  
 FT POLY-SER.  
 SQ SEQUENCE 300 AA; 33810 MM; 1199D44A28F7251 CRC64;  
 Query March 3, 7%; Score 128.5; DB 1; Length 300;  
 Best Local Similarity 26.4%; Pred. No. 0.084;  
 Matches 55; Conservative 24; Mismatches 66; Indels 49; Gaps 11;

DB 126 LPKXTE-----IQFAEN--TKPEVTKPVSEEE-----KXHYGV-RQR 162  
 QY 409 QHGRQARI--GRVAGNKDLYIGTFSQGEAAEYDVAAIKFRGANAVTNPDISRYVER 466  
 DB 163 PWGKAFAEIRDPNKKGRS-VWLGITPDIAAARAYDAAPFKGSKALINFPLE----- 215  
 QY 467 IMASSNLLAGELARRKQND-PRNKIDYNSVTSV-NNETVQVQAG-----NNNE 518  
 DB 216 -----VGKKKPRADGEKKRRDDDEKVTVEKVLKTEQSVAVNGGTFPPVSNLT 267  
 QY 519 NDEWKVLE-F-NHP 531  
 DB 268 ELCDMDLNGFLNFP 281  
 RESULT 11  
 ID RAD9 YEAST STANDARD; PRT; 1309 AA.  
 AC P14737; O04920;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE DNA repair protein RAD9.  
 GN RAD9 OR YDR217C OR YD9934.02C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_Taxid=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89313732; PubMed=2664461;  
 RA Schiestl R.H., Reynolds P., Prakash S., Prakash L.;  
 RT "Cloning and sequence analysis of the Saccharomyces cerevisiae RAD9  
 RT gene and further evidence that its product is required for cell cycle  
 RT arrest induced by DNA damage";  
 RL Mol. Cell. Biol. 9:1882-1896(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE=91061763; PubMed=2247073;  
 RA Weinert T.A., Hartwell L.H.;  
 RT "Characterization of RAD9 of Saccharomyces cerevisiae and evidence  
 RT that its function acts posttranscriptionally in cell cycle arrest after  
 RT DNA damage";  
 RL Mol. Cell. Biol. 10:6554-6564(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Murphy U., Harris D., Barrell B.G., Rajadream M.A.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP PHOSPHORYLATION, AND INTERACTION WITH RAD53.  
 RX MEDLINE=98429491; PubMed=9755168;  
 RA Valard J.E., Gilbert C.S., Green C.M., Lowndes N.F.;  
 RT "The budding yeast Rad9 checkpoint protein is subjected to  
 RT Mec1/Trf1-dependent hyperphosphorylation and interacts with Rad53  
 RT after DNA damage";  
 RL EMBO J. 17:5679-5688(1998).  
 CC -1- FUNCTION: Essential for cell cycle arrest at the G2 stage  
 CC following DNA damage by X-irradiation or inactivation of DNA  
 CC ligase.  
 CC -1- SUBUNIT: Physically associates with RAD53.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: Contains 1 BRCT domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M26049; AAA34954.1; -  
 DR EMBL; Z48612; CAA88497.1; -

DR FIR; S59424; BVBYD9.  
 DR PDB; 1FHR; 18-OCT-00.  
 DR PDB; 1J4K; 05-DEC-01.  
 DR PDB; 1J4L; 05-DEC-01.  
 DR PDB; 1J4P; 05-DEC-01.  
 DR PDB; 1J4Q; 05-DEC-01.  
 DR PDB; 1K2M; 05-DEC-01.  
 DR PDB; 1K2N; 05-DEC-01.  
 DR PDB; 1K3N; 05-DEC-01.  
 DR PDB; 1K3Q; 05-DEC-01.  
 DR Germline; 140709; .  
 DR SGD; S0002625; RAD9.  
 DR GO; GO:0005634; C:nucleus; IC.  
 DR GO; GO:0005615; P:protein binding; IDA.  
 DR GO; GO:0000077; P:DNA damage response; signal transduction re. . .; IMP.  
 DR GO; GO:0006289; P:nucleotide excision repair; IMP.  
 DR GO; GO:0045944; P:positive regulation of transcription from P. . .; IMP.  
 DR GO; GO:0000074; P:regulation of cell cycle; IGI.  
 DR InterPro; IPR001357; BRCT.  
 DR Pfam; PF00533; BRCT; 1.  
 DR SMART; SM00292; BRCT; 1.  
 DR PROSITE; PS00172; BRCT; 1.  
 DR Cell cycle; DNA damage; DNA replication inhibitor; Nuclear protein;  
 KW Phosphorylation; 3D-structure.  
 FT DOMAIN 994 1122 BRCT.  
 FT CONFLICT 433 433 C -> S (IN REF. 3).  
 SQ SEQUENCE 1309 AA; 148413 MW; 6B7D39A95021F64 CRC64;  
 Query Match 3.6%; Score 127.5; DB 1; Length 1309;  
 Best Local Similarity 18.6%; Pred No. 0.82; Indels 205; Gaps 32;  
 Matches 132; Conservative 103; Mismatches 269;  
 QY 37 VPTTFYSSQSHLSNFGMCYGVNGNENFHSPLTPYLPKSDGSLCILEALRSGQVWVP 96  
 DB 455 IKTIQINSPEQNAL-----NATFEFVTLIRINFEPILEVETSSPSKNTSKP 503  
 QY 97 TSS---PKLEDFLGATWGTHERYSGHERGLSDSYVNS--QNAEAQPRDL-----SQPF 148  
 DB 504 SNSPFPKED-----TFNHERVEVNNVFSNDIDNSSMAARDIILAGSDF 553  
 QY 149 RQGHNSVOTHPYSG--LACHGLYQ---APLEETTKETHVS---DCSGLMQMTGL 199  
 DB 554 NEKEKLTDRILYQLSGKQSDSGSETERMSFNEADTKETMSVEVLYDELPEVEEQ 613  
 QY 200 KMWAVPTREFSTHQVLEQONNCGMNERNGVSLGVCGLQSLISMSGSGSCVTA 259  
 DB 614 DLOTSFKKIVEEETLME--IKSKKN-----SLQTHDNEKNSDK 653  
 QY 260 PSQTSVAVDAKKGHAKLQKQPYHRSIDTF--GRTQYGVTRHRTGRYE--- 312  
 DB 654 QDQTES--LDVALIEHESKGSSEQKXLMQFPESEOEIIQNRRTIRKQDIEEGEE 711  
 QY 313 -----AHLMN-----SCKE-----GOTKRGROYLGGDMEKARA 346  
 DB 712 EENRSTKSTPTKHLKNSDLDAASIKRHSCTITQTEPESGKDSKSGSVPEE-GIRT 770  
 QY 347 YDLALK---YWGPESTHNSIENYQ---VQLEEMKMGROBYVAHLRKSXSGFRGA 398  
 DB 771 ADNSFLSKDILFGNAVWCQYT-WNYKFPYGLLEVDNOD-----GC 812  
 QY 399 SIY-----RGVTRHGHGRWQARIGRYA--GNKDLVLTGFTQEE-----AAAYDAAI 446  
 DB 813 WIFYEFGRLTQEDDIYVIDRIGAVTFDNGEYVVGLEGRSHNIIIRIRQYDVHL 872  
 QY 447 KFRGANAVTWFDISRYDVERIMASSNLAAGELARKK--DNDFPKQIDY----- 494  
 DB 873 KKNKASGL--LGRKTLIKALSSISLDSSEWAKAKIILDENKCKGDARVYRHPD 928  
 QY 495 NKGVV-----TSVNEETVQVQGNNNNNDEEMQVLFNHSQCOQANGSGDOKI 546  
 DB 929 RKSMTVLSRKHGTDDEKQINHTHEVYNNIESSSEKKEIV-KQDSRDLALFHAGASL 987  
 QY 547 MNCNTRNSAFSVALQDLIGDSVSGQHNLDESSKIGTFHSNTSLVLTSSREASP 606

DB 988 FSSGEIRT-----GNVFKCIFVLSTLEENRELR 1017  
 QY 607 E-----KRGSLPMP-PMEKIVNPIS-----VTSWLP 637  
 DB 1018 QTESQSGTVEBSGSLTFNFTHLPLASLVNKNKTDIRLALAKLP 1066  
 RESULT 12  
 ID ATRX\_PANTR STANDARD; PRT; 2492 AA.  
 AC ATRX\_PANTR  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP).  
 GN ATRX.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22763540; PubMed=12777533;  
 RA Kitano T., Schwarz C., Nickel B., Pabo S.;  
 RT "Gene diversity patterns at 10 X-chromosomal loci in humans and chimpanzees";  
 RL Mol. Biol. Evol. 20:1281-1289(2003).  
 CC -!- FUNCTION: Could be a global transcriptional regulator. Modifies gene expression by affecting chromatin. May be involved in brain development and facial morphogenesis.  
 CC -!- SUBUNIT: Probably binds EZH2. Binds annexin V in a calcium and phosphatidylcholine/phosphatidylserine-dependent manner (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear. Associated with pericentromeric heterochromatin during interphase and mitosis, probably by interacting with Hpi.  
 CC -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.  
 CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC EMBL; AB102642; BAC81111.1; .  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR000330; SNF2\_N.  
 DR Pfam; PF00271; Helicase\_C\_1.  
 DR Pfam; PF00176; SNF2\_N; T.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC\_C; 1.  
 KW DNA repair; Nuclear protein; DNA-binding; Helicase; ATP-binding; Zinc-finger.  
 KM  
 FT ZN FING 220 268  
 FT NP BIND 1594 1601  
 FT SITE 1719 1722  
 FT DOMAIN 745 750  
 FT DOMAIN 1151 1156  
 FT DOMAIN 1166 1169  
 FT DOMAIN 1202 1206  
 FT DOMAIN 1259 1266  
 FT DOMAIN 1443 1466  
 FT DOMAIN 1499 1502  
 FT DOMAIN 1929 1939  
 FT DOMAIN 1941 1948  
 FT DOMAIN 2232 2235  
 FT DOMAIN 2262 2265  
 FT POLY-GLU.  
 PHD-TYPE.  
 ATP (POTENTIAL).  
 DEGR BOX.  
 POLY-SER.  
 POLY-SER.  
 POLY-LYS.  
 POLY-ASP.  
 POLY-GLU.  
 POLY-LYS.  
 POLY-LYS.  
 POLY-SER.  
 POLY-LYS.







RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=91330307; PubMed=1678319;  
 RA Joshi G., Carlson M., Thilveris A., Albertsen H., Gelbert L.,  
 RA Samowitz W., Gorden J., Stevens J., Spirio L., Robertson M.,  
 RA Sargent L., Krapcho K., Wolff E., Butt R., Hughes J.P.,  
 RA Warrington J., McPherson J.D., Wamuth J., Le Fautier D.,  
 RA Abderrahim H., Cohen D., Leppert W., White R.,  
 RT "Identification of deletion mutations and three new genes at the  
 RT familial polyposis locus.";  
 RL Cell 66:601-613 (1991).  
 RN [3]  
 RP ASSOCIATION WITH CATENINS.  
 RX MEDLINE=94082295; PubMed=8259519;  
 RA Su L.-K., Vogelstein B., Kinzler K.W.,  
 RT "Association of the APC tumor suppressor protein with catenins.";  
 RL Science 262:1734-1737 (1993).  
 RN [4]  
 RP DISEASE.  
 RX MEDLINE=95174843; PubMed=7661930;  
 RA Hamilton S.R., Liu B., Parsons R.E., Papadopoulos N., Jen J.,  
 RA Powell S.M., Krush A.J., Berk T., Cohen Z., Tetu B., Burger P.C.,  
 RA Wood P.A., Taqi F., Bookier S.V., Petersen G.M., Offenhans G.J.A.,  
 RA Termette A.C., Giardiello F.M., Vogelstein B., Kinzler K.W.,  
 RT "The molecular basis of Turcot's syndrome.";  
 RL New Engl. J. Med. 332:839-847 (1995).  
 RN [5]  
 RP DISEASE.  
 RX MEDLINE=97094176; PubMed=8940264;  
 RA Eccles D.M., van der Luijt R.B., Breukel C., Bullman H., Bunyan D.,  
 RA Fisher A., Barber J., du Boulay C., Pimrose J., Burn J., Podde R.,  
 RT "Hereditary desmoid disease due to a frameshift mutation at codon 1924  
 RT of the APC gene.";  
 RL Am. J. Hum. Genet. 59:1193-1201 (1996).  
 RN [6]  
 RP DISEASE.  
 RX MEDLINE=20243021; PubMed=10782927;  
 RA Couture J., Mitrì A., Lagace R., Saults R., Berk T., Bouchard H.-L.,  
 RA Podde R., Altman B., Bapat B.,  
 RT "A germline mutation at the extreme 3-prime end of the APC gene  
 RT results in a severe desmoid phenotype and is associated with  
 RT overexpression of beta-catenin in the desmoid tumor.";  
 RL Clin. Genet. 57:205-212 (2000).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 2-55.  
 RX MEDLINE=20384842; PubMed=10926498;  
 RA Day C.L., Alber T.,  
 RT "Crystal structure of the amino-terminal coiled-coil domain of the  
 RT APC tumor suppressor.";  
 RL J. Mol. Biol. 301:147-156 (2000).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 1021-1035 IN COMPLEX WITH  
 RP CTNNB1.  
 RX MEDLINE=21564054; PubMed=11707392;  
 RA Eklof Spink K., Fridman S.G., Weis W.I.,  
 RT "Molecular mechanisms of beta-catenin recognition by adenomatous  
 RT polyposis coli revealed by the structure of an APC-beta-catenin  
 RT complex.";  
 RL EMBO J. 20:6203-6212 (2001).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 2034-2049 IN COMPLEX WITH  
 RP AXIN.  
 RX MEDLINE=20271867; PubMed=10811618;  
 RA Spink K.E., Polakis P., Weis W.I.,  
 RT "Structural basis of the axin-adenomatous polyposis coli  
 RT interaction.";  
 RL EMBO J. 19:2270-2279 (2000).  
 RN [10]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=94154728; PubMed=8111410;  
 RA Nagase H., Nakamura Y.,  
 RT "Mutations of the APC (adenomatous polyposis coli) gene.";  
 RL Hum. Mutat. 2:425-434 (1993).

RN [11]  
 RP VARIANTS FAP.  
 RX MEDLINE=91335211; PubMed=1651563;  
 RA Nishisho I., Nakamura Y., Miyoshi Y., Miki Y., Ando H., Horii A.,  
 RA Koyama K., Utsunomiya J., Baba S., Hedge P., Markham A., Krush A.J.,  
 RA Petersen G.M., Hamilton S.R., Nildert M.C., Levy D.B., Bryan T.M.,  
 RA Preisinger A.C., Smith K.J., Su L.-K., Kinzler K.W., Vogelstein B.,  
 RT "Mutations of chromosome 5q21 genes in FAP and colorectal cancer  
 RT patients.";  
 RL Science 253:665-669 (1991).  
 RN [12]  
 RP VARIANTS FAP.  
 RX MEDLINE=93265030; PubMed=1338904;  
 RA Miyoshi Y., Nagase H., Ando H., Ichii S., Nakatsuru S., Aoki T.,  
 RA Miki Y., Mori T., Nakamura Y.,  
 RT "Somatic mutations of the APC gene in colorectal tumors: mutation  
 RT cluster region in the APC gene.";  
 RL Hum. Mol. Genet. 1:229-233 (1992).  
 RN [13]  
 RP VARIANTS FAP.  
 RX MEDLINE=93244793; PubMed=1338691;  
 RA Nakatsuru S., Yanagisawa A., Ichii S., Tahara E., Kato Y.,  
 RA Nakamura Y., Horii A.,  
 RT "Somatic mutation of the APC gene in gastric cancer: frequent  
 RT mutations in very well differentiated adenocarcinoma and signet-ring  
 RT cell carcinoma.";  
 RL Hum. Mol. Genet. 1:559-563 (1992).  
 RN [14]  
 RP VARIANTS FAP TRP-1348, AND VARIANTS ASP-1118; MET-1292; VAL-1304 AND  
 RP SER-2502  
 RX MEDLINE=93250848; PubMed=1338764;  
 RA Nagase H., Miyoshi Y., Horii A., Aoki T., Petersen G.M.,  
 RA Vogelstein B., Maher E., Ogawa M., Maryama M., Utsunomiya J.,  
 RA Baba S., Nakamura Y.,  
 RT "Screening for germline mutations in familial adenomatous polyposis  
 RT patients: 61 new patients and a summary of 150 unrelated patients.";  
 RL Hum. Mutat. 1:467-473 (1992).  
 RN [15]  
 RP VARIANTS FAP TRP-99.  
 RC TISSUE=Peripheral blood lymphocytes;  
 RX MEDLINE=95134544; PubMed=7831149;  
 RA Dobbie Z., Spycher M., Huertlman R., Ammann R., Ammann T., Roth J.,  
 RA Mueller A., Mueller H., Scott R.J.,  
 RT "Mutation analysis of the first 14 exons of the adenomatous  
 RT polyposis coli (APC) gene.";  
 RL Eur. J. Cancer 30A:1709-1713 (1994).  
 RN [16]  
 RP VARIANTS FAP GLY-722.  
 RX MEDLINE=95135430; PubMed=7831931;  
 RA Stella A., Montero M., Resta N., Marchese C., Susca F., Gentile M.,  
 RA Romio L., Pilla S., Prete F., Varenti C., Guanti G.,  
 RT "Four novel mutations of the APC (adenomatous polyposis coli) gene in  
 RT FAP patients.";  
 RL Hum. Mol. Genet. 3:1687-1688 (1994).  
 RN [17]  
 RP ERRATUM.  
 RA Stella A., Montero M., Resta N., Marchese C., Susca F., Gentile M.,  
 RA Romio L., Pilla S., Prete F., Varenti C., Guanti G.,  
 RL Hum. Mol. Genet. 3:1918-1918 (1994).  
 RN [18]  
 RP VARIANTS FAP ILE-171.  
 RX MEDLINE=97144176; PubMed=8990002;  
 RA van der Luijt R.B., Meera Khan P., Vaseen H.F.A., Tops C.M.J.,  
 RA van Leeuwen-Cornelisse I.S.J., Wijnen J.T., van der Klift H.M.,  
 RA Plug R.J., Griffioen G., Podde R.,  
 RT "Molecular analysis of the APC gene in 105 Dutch kindreds with  
 RT familial adenomatous polyposis: 67 germline mutations identified by  
 RT DGGE, PTT, and Southern analysis.";  
 RL Hum. Mutat. 9:7-16 (1997).  
 RN [19]  
 RP VARIANTS COLORECTAL CARCINOMA THR-880; ILE-890 AND VAL-1508.  
 RX MEDLINE=98080146; PubMed=9419797;  
 RA Miyaki M., Nishio T., Konishi M., Kikuchi-Yanoshita R., Tanaka K.,

RA Muraoka M., Nagato M., Chong J.-M., Koike M., Terada T., Kawahara Y.,  
 RA Fukutome A., Tomiyama J., Chuganji Y., Momoi M., Utsunomiya J.,  
 RA "Drastic genetic instability of tumors and normal tissues in Turoc  
 RT syndrome.";  
 RL Oncogene 15:2877-2881(1997).  
 RN [20]  
 RP VARIANT LYS-1307.  
 RX MEDLINE=98400248; PubMed=9731522;  
 RA Redson M., Nathanson K.L., Yuan Z.Q., Neuhausen S.L., Saragopan J.,  
 RA Wong N., Yang D., Nafa D., Abrahamson J., Orzellik H.,  
 RA Aulin-Ozerkis D., Andrulis I., Daly M., Finsky L., Schrag D.,  
 RA Gallinger S., Kaback M., King M.-C., Woodage T., Brody L.C.,  
 RA Godwin A., Warner E., Weber B., Foulkes W., Offit K.;  
 RT "The APC 11307K allele and breast cancer risk.";  
 RL Nat. Genet. 20:13-14(1998).  
 RN [21]  
 RP VARIANTS LYS-1307 AND GLN-1317.  
 RC TISSUE=Peripheral blood;  
 RX MEDLINE=98393712; PubMed=9724771;  
 RA Frayling I.M., Beck N.E., Ilyas M., Dove-Edwin I., Goodman P.,  
 RA Pack K., Bell J.A., Williams C.B., Hodgson S.V., Thomas H.J.W.,  
 RA Talbot I.C., Bodmer W.F., Tomlinson I.P.M.;  
 RT "The APC variants 11307K and 11317Q are associated with colorectal  
 RT tumors, but not always with a family history.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10722-10727(1998).  
 RN [22]  
 RP VARIANT LYS-1307.  
 RX MEDLINE=98400259; PubMed=9731533;  
 RA Woodage T., King S.W., Macholder S., Hartge P., Struwing J.P.,  
 RA McAdams M., Laken S.J., Tucker M.A., Brody L.C.;  
 RT "The APC 11307K allele and cancer risk in a community-based study of  
 RT Ashkenazi Jews.";  
 RL Nat. Genet. 20:62-65(1998).

Query March 3.5%; Score 121; DB 1; Length 2843;

Best Local Similarity 19.6%; Pred. No. 7.1;  
 Matches 131; Conservative 97; Mismatches 268; Indels 174; Gaps 30;

QY 58 GVGNGNFHSPVTW--PLKSDGSLCTIEALKRSQTVMT--SSPKLEDFLGATMGTH 114  
 DB 2235 GVRNSSSTSPVSKGPKLTPAS---KSPSEQITATSPRGAKPSVKSELSFVARQTS 2290  
 QY 115 EYGSHERGLSDSTIYVNSQNAEAPNRDLISQPRQCHMSVQTHPYSGLA-CHGLYQA 173  
 DB 2291 QIGSSSKAPSRSG--SRDSTPSRPAQPLSRPISQSPRNSIS--PGRNGISPPNKLSQL 2345  
 QY 174 PLEETTRKTHVSDSSLMPTQTEGLKNVAPTEPSTHQVLEQVNC-----222  
 DB 2346 P---RTSSPSTASTKS-----GSGKMSYTPGRQMS--QONTLTKQTGLSKNMSIPRSE 2395  
 QY 223 ---GNGERNGVSLISGVCGLQSLISMSPGSSSCVTPASGTDVAVDKKRGHAKL 278  
 DB 2396 SASKGLNQMNNG--NGANKVELSNMSTKSGSESDRSERP-----VIVRGSTPIKE 2446  
 QY 279 GQKQPVHRKSIDTFQRTSQYRGVTRHRWGRYEAMHMDNSCKKEGQTRKGRQVYLGCD 338  
 DB 2447 APSPTLRKRL-----EEGASFSLS-----PSSRPASPTRSQAQTPVLSPS 2487  
 QY 339 MEKKAARAYDLAALKYWGSPSTHINSINYQ-----VQLEMKMSRQEVVAHLRR 389  
 DB 2488 LPDMGL-----STHSSVQAGWRLPPNLSPTIEYNDGRPAKHD---IAR 2530  
 QY 390 KSGGFSRGASIRYGVTRHGHGRWQARIGRA-----GNKDLVLTGTFSTOEAAEAYD-- 442  
 DB 2531 SHSEBPSRLPIRSGTWKREHSHSSSLPRVSTWRTGSSSILSASSESEKASEDEK 2590  
 QY 443 ---VAATFRGANAYTNFDIRYVERIMASSNLLAGELARKKNDPRNKDIDYNSV 498  
 DB 2591 HYNISIGTKOSKENOVSAKGTWRKI KENEFSPTN-STSQTVSSGATNGAESKTLIYOMAP 2649  
 QY 499 VTS-----VNEETVOVQAGN-----NNNENDESMRWVLFNHPISQ 533  
 DB 2650 AVSKTEDVWVRLEDGCIINNPRGSRGSPGTGTPPVIDSVSEKANPNIKDSK-----DNO 2701

QY 534 QOQANGNS-DOKIMGNRYNSAFSVALQDLIGIDSVSGGQHNMLEDSSKIGTHFNTS 592  
 DB 2702 AKQVNGSVPMRTVGLNRLNSFIQVADPDQKQTE-IPKQNNPVVSB-----TNES 2754  
 QY 593 SLV-----TSSSRERASPEKRGPSLLFPMPPMETK-VNPIGTSTV--SWLPSPTVQMRP 645  
 DB 2755 SIVERTPPSSSSSKHSSPS-----GTYAARVTFPNINPSF-----RK 2792  
 QY 646 SPAISLSHP 655  
 DB 2793 SSADSTSRP 2802

Search completed: March 9, 2004, 10:46:04  
 Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:41:58 ; Search time 45 Seconds  
(without alignments)

4648.635 Million cell updates/sec

Title: US-10-024-632-2

Sequence: 1 MKRINSSNTDGNHNNWLG.....RSPALSLHLPVPSWTD 663

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL.25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacterioph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1370.5	39.1	639	10	Q84Z02
2	1336	38.1	588	10	Q85W7
3	1328	37.9	555	10	Q88914
4	1328	37.9	555	10	Q42462
5	956	27.3	692	10	Q881E1
6	953	27.2	584	10	Q8LSN1
7	947	26.9	579	10	Q8LSN2
8	942	26.9	579	10	Q8LSN3
9	923.5	26.4	425	10	Q9CAH3
10	919.5	26.3	597	10	Q8LGG3
11	918.5	26.2	555	10	Q7XPD4
12	916.5	26.1	581	10	Q7LFR2
13	900.5	25.7	555	10	Q9F1E2
14	893.5	25.5	540	10	Q9L737
15	890	25.4	485	10	Q41832
16	879	25.1	516	10	Q9SYC2

17	841.5	24.0	492	10	Q8LGG4	Q8LGG4	oryza sativ
18	839.5	24.0	437	10	Q9LSM4	Q9LSM4	arabidopsis
19	831	23.7	566	10	Q9LXAX	Q9LXAX	arabidopsis
20	688.5	19.7	498	10	Q7XTE2	Q7XTE2	oryza sativ
21	672	19.2	345	10	Q9AN44	Q9AN44	arabidopsis
22	645	18.4	332	10	Q9S9M9	Q9S9M9	arabidopsis
23	616	17.6	308	10	Q9MAD6	Q9MAD6	arabidopsis
24	566	16.2	423	10	Q8RV98	Q8RV98	oryza sativ
25	534	15.2	266	10	Q7XPM5	Q7XPM5	oryza sativ
26	529	15.1	199	10	Q8GXZ5	Q8GXZ5	arabidopsis
27	517	14.8	205	10	Q9M2V4	Q9M2V4	arabidopsis
28	515.5	14.7	533	10	Q9AX14	Q9AX14	pisum sativ
29	498.5	14.2	505	10	Q84U24	Q84U24	antirrhinum
30	497	14.2	519	10	Q9XHD4	Q9XHD4	petunia hyb
31	496	14.2	423	10	Q22949	Q22949	arabidopsis
32	495.5	14.1	505	10	Q84U23	Q84U23	antirrhinum
33	495	14.1	423	10	Q8LA00	Q8LA00	arabidopsis
34	487	13.9	372	10	Q84TV8	Q84TV8	oryza sativ
35	487	13.9	434	10	Q84TB5	Q84TB5	oryza sativ
36	484	13.8	549	10	Q8VWM8	Q8VWM8	malus domes
37	483.5	13.8	428	10	Q8GWR2	Q8GWR2	arabidopsis
38	483	13.8	446	10	P93632	P93632	zea mays (m
39	477.5	13.6	436	10	Q8H443	Q8H443	oryza sativ
40	477	13.6	440	10	Q8VWQ6	Q8VWQ6	hordeum vul
41	475.5	13.6	352	10	Q9F9H5	Q9F9H5	arabidopsis
42	474.5	13.5	633	10	Q9FV33	Q9FV33	plcea abies
43	473.5	13.5	531	10	Q9FV34	Q9FV34	plcea abies
44	467.5	13.3	433	10	Q65008	Q65008	zea mays (m
45	462	13.2	457	10	Q9XHD3	Q9XHD3	petunia hyb

## ALIGNMENTS

## RESULT 1

ID Q84Z02 PRELIMINARY; PRT; 639 AA.

AC Q84Z02; 01-JUN-2003 (TREMBLrel. 24, Created)  
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Putative ovule development protein antilegumena (ANT).  
GN P0474G09.16.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzoideae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RF SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Katayose Y.,  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC  
clone: P0474G09.";  
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: APO05309; BAC56815.1; -  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro: IPR001471; TF\_ERF.  
DR Pfam: PF00847; AP2-domain; 2.  
DR PRINTS: PR00367; ETRSPLEMMT.  
DR ProDom: PD001423; TF\_ERF; 2.  
DR SMART: SM00380; AP2; 2.  
DR SWARP: SM00380; AP2; 2.  
SQ  
SEQUENCE 639 AA; 67622 NM; EB4BC5123C56568F CR664;

Query Match 39.1%; Score 1370.5; DB 10; Length 639;

Best Local Similarity 46.6%; Pred. No. 1.6e-90;

Matches 337; Conservative 68; Mismatches 169; Indels 149; Gaps 30;

QY 4 INSSNTDGN---HNWIGFSLPMKWEATSAATVPTTFW-----SPGSH-- 49  
DB 1 MINSNNGNGTVAASGWLFSLSFPM-----ASTMDHHVHHHQQQQQQQQQQHHQ 55

```

QY 50 -----LSNFGWCY-----VGENGNFHSPLTVMPLKSDGSLICEALKRSQTO 92
DB 56 QOQHLLFPSPVTTAAAAAAYGAGVVAATNGYSQASLMSPLKSDGSLICEALRRDPOD 115
QY 93 VMTPTSSPLDEDFLGATMGTHYSGHERGSLD---STYNSQNAEAQPRDLISQPR 149
DB 116 ---HHGPRLEDFLGAA-----QPMALSLDNTSSFYGGGGAAGH-----G 155
QY 150 QOQHNSVOTHPYSGLAGHGYOAPLEBETKETHVSCSSSLMPOMTEGLKNVAPTRRF 209
DB 156 QHGYL--QACDLYGPPAPSLVTADEAAA-----AAAAASWVAAGAA 199
QY 210 STHQOVLBQNNCGMNGERNNGVLSGVCCEL--QSLSNSPGSQSCVTAAPS--GTDVA 267
DB 200 TAY-----ATGAADANAENVLPBATAOHLHPLALSNSSGSLSCITAGEYGMAYVA 253
QY 268 V--DAKRGHA-KLGOKOPV--HRKSIDTFGORTSGYRGVTRHRTGRYEALHMDNSCKE 323
DB 254 AADGGRKGGAGGGGQOKOPVHHRKSIDTFGORTSGYRGVTRHRTGRYEALHMDNSCKE 313
QY 324 GQTRKGRQVYLGGYDMEERKAAAYDLAALKYWGSPSTHINFSIENYOVQLEBKMSRQCY 383
DB 314 GQTRKGRQ---GGYDMEERKAAAYDLAALKYWGSPSTHINFLDYQELBEMKMTROCY 370
QY 384 VAHLRRKSSGFRGASITRGVTRHGHQGRWQARIGRVAGNCDLYLGTSTQEBAAEAYDV 443
DB 371 VAHLRRKSSGFRGASITRGVTRHGHQGRWQARIGRVAGNCDLYLGTSTQEBAAEAYDV 430
QY 444 AAIKRGANAATNPFISRVDYERIMASSNLAGEIARRKNDPBNKDIDYKNSVTVSVN 503
DB 431 AAIKRGANAATNPFISRVDYERIMASSNLAGEIARRKNDPBNKDIDYKNSVTVSVN 477
QY 504 NEET-----VOYQAKNNNNENDSEKMYLFTNHPGQQOQANG---NSDQKIMCGMYR 553
DB 478 DHSAGAGAIIVSHSAADIAVADTLMKAT--TAPRQOQOHHDDVVLGADQ----- 526
QY 554 NSAFSMALODLIGIDVSGQGNMLDESEKIGTHFSNTSLVTSLSRREASPEKRGD-S 612
DB 527 -AATSV-LHDLVAVDAAAHAHQOQOQAHSM---SAASLVTLSLSNRREGSPRGGSL 580
QY 613 LLEPMP-----PMETKIVN--PTGISTVWLPSPYQMPSPALISLHPVVAS 659
DB 581 MFAKSPAPVAAASAOQAATKLAAPLPLG---SWVSPSPASARP--PAVSIAHVPPLPA 635
QY 660 WTD 662
DB 636 WTD 638

```

## RESULT 2

```

Q885W7 PRELIMINARY; PRT; 588 AA.
ID Q885W7
AC Q885W7;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DE Putative ovule development protein antilegumenta (ANT).
GN CUI175C11.21 OR CUI1017C11.11.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RT "Rice Genomic Sequence".
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Wang R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Currie J.,

```

```

RA Collura K.;
RT "Rice Genomic Sequence."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DB EMBL; AC103891; AAM19141.1;
DB EMBL; AC135157; AAC00690.1;
DB Gramene; Q885W7;
DR GO; GO:0005634; Cytosol; IEA.
DR GO; GO:0003700; Fertilization factor activity; IEA.
DR GO; GO:0006355; Regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PRO0367; ETHRSELEMENT.
DR Prodom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2; 2.
DR SMART; SM00380; AP2; 2.
SQ SEQUENCE 588 AA; 63230 KM; A13087F5A36CEDIC CR664;

```

```

Query Match 38.1%; Score 1336; DB 10; Length 588;
Best Local Similarity 47.5%; Pred. No. 4; 3e-88;
Matches 333; Conservative 62; Mismatches 146; Indels 160; Gaps 30;

```

```

QY 7 SNTDGNHNNMLGFLSHHMKKATSAATVPTTFMSPSQSHLSNFGWCYGVGN--GN 64
DB 2 SPPDATTGNTL-FS-PPAQVAPV---PGYIV-----GAYGDTSTAV 45
QY 65 FHSPLTVMPLKSDGSLICEALKRSQTVMPPTSPKEDPLGATMGTHYSGHERGSL 124
DB 46 YSHLPVMTLKSDGSLICMEG-----MMPSSPKEDPLGCGNGSGH----- 87
QY 125 LDSITYNSQNAEAQPRDLISOPFRQGMNSQTHYHYSGLACHGYOAPLEBETKETH 184
DB 88 -DPATYSGQOEED-----ASRAAYQHQLV--PYN-----YQPLTEMLQEAA 130
QY 185 VSDCSLMPOMTEGLKNVAPTRFEFTHQOVLBQNNCGMNGERNNGVLSGVCCEL--QS 243
DB 131 AAMPEDAM-----AAKNF-----LVTSAGAYGNQ-----EMQP 161
QY 244 LSLNSPSPGQ--SSCTV--APSGDTSVAVDA-----KRGAKIG 279
DB 162 LSLNSPSPGQSSSCVSAAPQOQHQAIVAAAAAGDQGSNSNDGEQVGRKRGKGG 221
QY 280 QKOPVHRKSIDTFGORTSGYRGVTRHRTGRYEALHMDNSCKEQTGRQVYLGGYDM 339
DB 222 QKOPVHRKSIDTFGORTSGYRGVTRHRTGRYEALHMDNSCKEQTGRQVYLGGYDM 278
QY 340 EERAAAYDLAALKYWGSPSTHINFSIENYOVQLEBKMSRQCYVAHLRRKSSGFRGAS 399
DB 279 EERAAAYDLAALKYWGSPSTHINFSIENYOVQLEBKMSRQCYVAHLRRKSSGFRGAS 338
QY 400 IYRGVTRHGHQGRWQARIGRVAGNCDLYLGTSTQEBAAEAYVAAIKRGANAATNPF 459
DB 339 IYRGVTRHGHQGRWQARIGRVAGNCDLYLGTSTQEBAAEAYVAAIKRGANAATNPF 388
QY 460 SRVDYERIMASSNLAGEIARRK-----DNDPRKNDIDYKNSVTVSVNNEETVOAG 513
DB 399 TRVDYERIMASSNLAGEIARRK-----DNDPRKNDIDYKNSVTVSVNNEETVOAG 450
QY 514 NNNNNENDSEKMYLFTNHPGQQOQANG---NSDQKIMCGMYRNSAFMALODLIGIDS 569
DB 451 -----TDMRWVL--HGSQOQOAACTEATADLQK---GWMGDAHS--ALHGIGFV 495
QY 570 VSGQGNMLDESEKIGTHFSNTSLVTSLSRREASPEKRGSLLEPMPMETKIVNP 628
DB 496 EESAADIDVPGKIGSINSSSLVTSLSNRREGSPRLGLAMLYAKH-----HPT 548
QY 629 GTSVTS---WLPSTVO---MRSPALISLHPVPASTWD 662
DB 549 AVSLAANNPMPMPAPAPAAHVPMPSPAL--AHLVPAAWTD 587

```

## RESULT 3

```

Q88914 PRELIMINARY; PRT; 555 AA.
ID Q88914
AC Q88914;

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE CAC.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 CX [1]  
 RN SEQUENCE FROM N.A.  
 RP Verciani P., Morandini P., Soave C.;  
 RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 DB EMBL, U44028; AAA86281.1; -  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR PRINTS; PR00167; ETRSPLENT.  
 DR ProDom; PD001623; TF\_ERF\_2.  
 DR SMART; SM00380; AP2; 2.  
 SO SEQUENCE 555 AA; 61741 MW; A627SD2F082B562D CRC64;

Query Match 37.9%; Score 1328; DB 10; Length 555;  
 Best Local Similarity 43.5%; Pred. No. 1.5e-87;  
 Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

1 MKRINESNTDGNHNLGFSLSPH -MKM-----EATSAATPTTFY 43  
 1 MKSPCDNDNHSNTNLGFSLSNNMKMGREGREATYSSSTCSAATSSSVPPQULV 60  
 44 SPSQHLNFGMCYCVGNGNFHSPFLTYMPLKSDGSLCEALKRSQTVVPTSPKLE 103  
 61 G---DNTNFGVCYSNNNGIYSHMSYMLRSDSLCLMALNRSSSHNHQSSPKVE 117  
 104 DFLGATMGTHYSGHERG--LSLDSITYNSQNAAPNRLLOPF---RQGHMSVQ 157  
 118 DFFG---THHNTTSHKAMDLSLDFYNTTH--EENTTNFOEFPSPQTNH--- 166  
 158 THPYSGLAGHGLYAPLEETETKETHVSDCSLMPQTEGLKNVAPTRFSTHQYLE 217  
 167 -----EEET--RNYGND-----PSLTHG----- 182  
 218 QQNMCGMNERGVSLGVGCEL-QSLSLSMSPGSSCVTAPS----- 261  
 183 -----GSFNVGV-----YGEFOQSLISMSPGSSCITGSHHQQONONHQSNH 229  
 262 -----GTVSAVDAKRGH---AKLGQKQPVHRKSIDTFGQRTSOYRGVTR 304  
 230 QQISEALVETSVGFETTTAAAKKRGQEDVVVQKQIVHRKSIDTFGQRTSOYRGVTR 269  
 305 HRMTGYEAHLMDNSCKEGQTRKROYVLLGYDMEBKAAAYDLAALKYGPSTHINS 364  
 299 HMTGYEAHLMDNSCKEGQTRKROYVLLGYDMEBKAAAYDLAALKYGPSTHINS 349  
 365 IENYQVQLEEMQNSROEVVAHLRKSQSGSFGATYGVYRHHQGMORIGVGNK 424  
 350 AANYQVEIEMQNSROEVVAHLRKSQSGSFGATYGVYRHHQGMORIGVGNK 409  
 425 DYLIGFTQEEAAEAAYDAALIKFGANAVTNFDISRYDVERIVASNLGELARKKD 484  
 410 DYLIGFTQEEAAEAAYDAALIKFGANAVTNFDITRDVDVRIMSSNTLLSGELARR-- 467  
 485 NDPRKNDIDYKTSVTVSNNESTVOYQAGANNNEDESEKVLPHNPSQQAANNGSDQ 544  
 468 -----NNSIV--VNTT-----DQ 479  
 545 KIMNGCNVNSAASMALQDLIGDSVSGQHMLDESSKIGTHFNTSLVTSLSRREA 604  
 480 TALNA-----VEGGSNNKV 494  
 605 SPEKRGSL--LEPMPMETKIV-NDIGTSVTSMLEPSTVQNRSPALSLSLHLPVFAWT 661

DB 495 STRPERLSPALFALPYQVQKMGFSGMGNMNSPMTSNFNAELK-TVALLTPQMPVFAMWA 553  
 CY 662 DT 663  
 DB 554 DS 555

RESULT 4  
 ID Q42462 PRELIMINARY; PRT; 555 AA.  
 AC Q42462;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE ANT (Putative ovule development protein antequemata) (OVULE  
 development protein ANTQUEMATA)  
 GN ANTQUEMATA OR T8819.30 OR AT4G37750.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 CX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=LANSBERG ERECTA, TISSUE=flower, and Ovule;  
 RC Elliott R.C., Betzner A.S., Huttner E., Oakes W.P., Tucker W.Q.J.,  
 RA Gerentes D., Perez P., Smyth D.R.;  
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA Klueber K., Chow H., Reiser L., Fischer R.L.;  
 RL Plant Cell 8:0-0(0).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP Bevan M., Van Der Schuren J., Chuang Y.-D., Voet M., Robben J.,  
 RA Volckaert G., Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP EU Arabidopsis sequencing project;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RP Robben J., Grymoprez B., Volckaert G., Mewes H.W., Lemcke K.,  
 RA Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RP EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RP Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.U., Tortum M., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.U., Sakurai T., Satou M.,  
 RA Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Arabidopsis Full Length cDNA Clones";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RN SEQUENCE FROM N.A.  
 RP Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.U.,  
 RA Tang C., Tortum M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.U., Sakurai T., Satou M.,  
 RA Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Arabidopsis Open Reading Frame (ORF) Clones";

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: U41339; AAB17364.1; -  
 DR EMBL: U40256; AAA91040.1; -  
 DR EMBL: AL035709; CAB38923.1; -  
 DR EMBL: AL161592; CAB90440.1; -  
 DR EMBL: AY080706; AAB50204.1; -  
 DR EMBL: AY117207; AAB51282.1; -  
 DR PIR: S71365; S71365.  
 DR TRANSFAC: T02639; -  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003700; P:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRSPPELEMT.  
 DR Prodom: PD001423; TF\_ERF; 2.  
 DR SEQUENCE 555 AA; 61725 MW; BD275D2F082B519B CRC64;

Query Match 37.9%; Score 1328; DB 10; Length 555;  
 Best Local Similarity 43.5%; Pred. No. 1.5e-87;  
 Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

QY 1 MKRINESNTDGGNNHMLGFSLSPH-MKM-----EATSAATVPTTFM 43  
 DB 1 MKSFCDNDNNHSTNTNLGFSLSNNMKMGGRGREALYSSSTSSAATSSSVPLV 60  
 QY 44 SPSGSHLSNFGMCYGVGENGFHSPDTMPLKSDGSLCILEALKRSQOVWVPTSSPKLE 103  
 DB 61 G---DNTSNFGVCYSNPNNGIYSHMSVMPFSDSLMLALNRSHSHNHQSSKVE 117  
 QY 104 DELGATWGTGTHYSGHSG--LSLDSIYNSQNAEAPNRDLISQPF---RQGGHNSVQ 157  
 DB 118 DFFG---THNNHSHKAMPLSDLSLFTYTH--BNPTTNNQEFSPFQTNH---- 166  
 QY 158 THPIYSGLAGHGLYAPLEETTKETHVSDCSLMPQTEGLKXWVAPTRFFSHQVLE 217  
 DB 167 -----EET--RNYND-----PSLTHG----- 182  
 QY 218 QQNMCGMNERNGVLSGVGCEL-QSLSLMSPGSSQSSCVTABS----- 261  
 DB 183 -----GSFNVG-----YGFQQSLSMSPGSSQSSCITGSHHQQNQNQHSGNH 229  
 QY 262 -----GTDVAVDAKRGH---AKLGQKQPVYRKSIDTFGQRTSQYRGVTR 304  
 DB 230 QOISEALVETSVGFETTMAAAKKRGQEDVYVVGQKQIVHRKSIDTGQRTSQYRGVTR 289  
 QY 305 HMTGTYEAHLMDNSCKEKGOTRGROYVLLGGYMEEKAAAYDAAALKYGPSTHINFS 364  
 DB 290 HMTGTYEAHLMDNSCKEKGOTRGROYVLLGGYMEEKAAAYDAAALKYGPSTHINFS 349  
 QY 365 IENYQVLEEMKMSRQEVVAHLRKSGSGSRGASITRGVTRHHQGMQARIGVAGNK 424  
 DB 350 AENYQKEIKIDMKMTQEVVAHLRKSGSGSRGASITRGVTRHHQGMQARIGVAGNK 409  
 QY 425 DLYIGFSTQOEAAEAYDVAATKFRGANAVTFDISYDVERIMASSNLLAGELARRKD 484  
 DB 410 DLYIGFSTQOEAAEAYDVAATKFRGTNAVTFDITRIDVDRIMSSNLLSELARRR-- 467  
 QY 485 NDPKNDIDYKSVTVSNNEETVQVQAGNNNNENDESEKVVLFNHPQQOQOQANGSDQ 544  
 DB 468 -----NNSIV--VANTE-----DQ 479  
 QY 545 KIMNCGYRNSAASMALDPLIGDSVSGQHMMLDESSKIGTHFSNTSSLYTSLSSREA 604  
 DB 480 TALNA-----VEGGSNKEV 494  
 QY 605 SPEKRGPSL--LFPMPPEYTKIV-NDIGTSVTSWLPSPVQKPPPAISLSHLVYFASVT 661  
 DB 495 STERELISPAIPALPQVQVQKMGSMGMSGMPWTSNPAELK--TVALTLTPQMPVFAAWA 553  
 QY 662 DT 663  
 DB 554 DS 555

## RESULT 5

Q8S1E1 PRELIMINARY; PRT; 692 AA.  
 ID Q8S1E1;  
 AC Q8S1E1;  
 DT 01-JUN-2002 (TRENBLREL. 21, Created)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)  
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)  
 DE Putative oule development protein aintegumenta-like protein.  
 GN P0035F12.3.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 RT clone: P0035F12."  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003313; BAB8946.1; -  
 DR Gramene; Q8S1E1; -  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003700; P:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRSPPELEMT.  
 DR Prodom: PD001423; TF\_ERF; 2.  
 DR SWART: SM00380; AP2; 2.  
 DR SEQUENCE 692 AA; 71515 MW; 4D5A0B49ED8772AF CRC64;

Query Match 27.3%; Score 956; DB 10; Length 692;  
 Best Local Similarity 35.3%; Pred. No. 1.7e-60;  
 Matches 263; Conservative 100; Mismatches 229; Indels 152; Gaps 28;

QY 16 HNWLGFSLSPHMKMEAT-----SAATVPTTFYSPSSSHLSNFGMCYGVGENGFHSP 68  
 DB 5 NNMILAPSLSPDQLPSSQNTSLISMAATTT-----AGDSSTDVCGNTPQDW----- 53  
 QY 69 LTMPLKSDGSLCILEALKRSQOVWVPTSSPKLEDPLGATMG--THEYSGH----- 119  
 DB 54 -----SMGSELSALY--AEKRLDFLGIGSFSSQGHNGKGVIPES 95  
 QY 120 -----ERGLSDSIYNSQNAEAPNRDL--SQPFQGGHMSVQTHPIYSGLAGHGL 170  
 DB 96 AACYASGSSSVGYLYPPSSSSSLQFADSVWVATSSP-----VVAH---DVSQGGM 144  
 QY 171 YQAPLEETTKETHVSDCSLMPQTEGLKXWVAPTRFFSHQVLEQQM--CGMGENB 229  
 DB 145 VSAIAAASGNGGIG-----LSMTKNMR--SQAPQAPQALSLSMNAAGTTAGG 194  
 QY 230 GVALGSV-GGCEL-----QSLSLMSPGSSQSSCVTA-----PSCGDS--VAVDAYK 272  
 DB 195 GGMALLAGAGRGRTTPASBSISTSAHGAATTATMAGKKEINEEGSGSAGAVVAAGSES 254  
 QY 273 RGH---AKLGQKQPVYRKSIDTFGQRTSQYRGVTRHMTGTYEAHLMDNSCKEKGOTRG 329  
 DB 255 GSGGAVVEAAGAAAARAKSVDTFGQRTSIYRGVTRKRWGRYEAHLMDNSCKEKGOTRG 314  
 QY 330 RQYVLLGGYMEEKAAAYDVAATKFGPSTHINFSIENYQVLEEMKMSRQEVVAHLNR 369  
 DB 315 RQ---GGYDEEKAAYDVAATKFGPSTTTTFPVNNEKELEEMKMTROGFVASSLR 371  
 QY 390 KSGGFSRGASITRGVTRHHQGMQARIGVAGNKDLYIGTSTQOEAAEAYDVAATKFR 449  
 DB 372 KSGGFSRGASITRGVTRHHQGMQARIGVAGNKDLYIGTSTQOEAAEAYDVAATKFR 431  
 QY 450 GANAVTFDISYDVERIMASSNLLAGELARRKNDPRNK-DIDYKSVTVSNNEETV 508



DB 432 GLNAVTFMSRYDKSLTDSALFVGTAAKRLDAEAAAYDVG---RIASHIGDGA 488  
QY 509 QVQAGNNNNNDSEWKLFFN-----HP-----SQQQQ-----ANGG 541  
DB 489 AAHYHHHHHAAAAAAPTIAFQAAAAPPHAAGLVHPYAPLPGWCKQEDHAIYAAHSL 548  
QY 542 SDQKIMNGC--NYRNSAFSMALQDILGIDSV-----GSGQNNML--DESSKIGTHFS 589  
DB 549 QDLHHLNIGAAAAAADFFSQMQQGHGIGSIDNASLEHSTGNSVTVNGDNGGGGGGYIM 608  
QY 590 NTSLSVTSLSRSREASPEKRGPSLLFPMPMETKIV--NPIGTSVTSMLSPPTVQMRSP 647  
DB 609 APMASVATATATVASHSHDGGCGKQVGMGYDSYLVGADAVGGGAGMPSMAMTPASAP 668  
QY 648 AITSLSH-----LPVFASTWDT 663  
DB 669 AATSSSDMTGVCHGAQLFSPVNDT 692

## RESULT 6

Q8LSN1 PRELIMINARY; PRT; 584 AA.  
AC Q8LSN1;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE BABY BOOM.  
GN BBW.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C24;  
RA Boutilier K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,  
RA Ouellet T., Zhang L., Hattori J., Liu C.-M., Miki B.L.A.,  
RA Custers J.B.M., van Lookeren Campagne M.M.;  
RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a  
RT conversion from vegetative to embryonic growth."  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
EMBL: AF317907; AAM33803.1;  
DR GO:0005634; C:nucleus; IEA.  
DR GO:0003700; F:transcription factor activity; IEA.  
DR GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro: IPR001471; TF\_ERF.  
DR Pfam: PF00847; AP2-domain; 2.  
DR PRINTS: PR00367; ETHRSPELEMT.  
DR PRODOM: PD001423; TF\_ERF; 2.  
DR SMART: SM00380; AP2; 2.  
SQ SEQUENCE 584 AA; 64515 MW; 1EA3DCDFIC900FB8 CRC64;

Query Match 27.2%; Score 953; DB 10; Length 584;  
Best Local Similarity 36.2%; Pred. No. 2, 2e-60;  
Matches 260; Conservative 75; Mismatches 174; Indels 210; Gaps 24;

DB 16 HNMVGLSPHMKMEATSAATVPTTFYMSQSLSNFGMCYGVGANGNHSPLTFLMK 75  
QY 5 NNMVGLSPH-----DONH-----HRTDVSSSTRITAVD 34  
DB 76 SDGSLCI-LEALKRSQTVWVPTSP---KLEDPLGATMGTHYSGHERGLSLDIYNN 131  
QY 35 VAGGYFDLAAPSDSSAVQTSFLSPGYTLERF-----TRDNNSHRMDIT- 81  
DB 132 SQNAEAQPNRDLISQPFROGSHNSVQTHPYSGLACHGLYQAPLEETTKETHVDCSSL 191  
QY 82 -----NGACNLT-----TNN- 93  
DB 192 MPQNTBGLKMWVAPTRP-PSYHQVLEQQMNCMGHERGVSIG-----SYGCG 239  
QY 94 -QNGPRLNFIQRTTTTITVNETVVDGNGDCGGGGGGGSGSLGSLMTKTMLSNHSVANA 151

QY 240 ELQ-----SLISMSBGSQSSCVTAPSGTDSVADAKKRGHAKLGQKQPVHKSIDTFGQ 294  
DB 152 NHQDNGNGARGSLSLMSNSTSDSNNTNNNDV---VQEKTIYDVETTP--KTTLESQ 206  
QY 295 RTSQYRGVTRHMTGRYEALWMDNSCKKEGQTRKQVYLLGGYMEKRAAYDIALAKY 354  
DB 207 RTSIYRGVTRHMTGRYEALWMDNSCKKEGQTRKQVYLLGGYMEKRAAYDIALAKY 266  
QY 355 WQSTHINSIENVQVLEKCMKSRQEVYALRRKSGFSGASITYGVTRHHQGRQ 414  
DB 267 WQSTHINSIENVQVLEKCMKSRQEVYALRRKSGFSGASITYGVTRHHQGRQ 326  
QY 415 ARIGRVAGKNDLYLGTFTQEEAAEYDVAIKFGANAVTFDISRYDVERIMASSNL 474  
DB 327 ARIGRVAGKNDLYLGTFTQEEAAEYDVAIKFGANAVTFDMRVYVKAILESPLP 386  
QY 475 AGEIARKKQNDPRKNDIDYNSVTVSNNETVQVQGN--NNNENDEWKLFFNHFS 532  
DB 387 IGSSAKRLD-----VNNPVAMNINNVSEANNSVGMONTAFQHQ 429  
QY 533 -----QCCQANGSDQKIMNCGNYRNSAFSMALQDILGIDSVGSGQNNMLDESSKIG 585  
DB 430 GMDLSLQQQQERYV-----YNGGSLSTESTRVCX-----QEEQGHFLNPSHMT 479  
QY 586 T-HFSNTSLVTSLSRSREASPEKRGPSLLFPMPMETKIVNPIGTSVTSMLSPPTV-- 641  
DB 480 NYDHSSTDDSVTVGANGVSYGYGQFAI-----PVGTSV-NDPFTAAEI 525  
QY 642 -----QMRPSP-----AITSLSH-----LPVFASTWDT 663  
DB 526 ANNAHHYVYAAHQQQQQQIQSBGDFPVAISNNHSMNYFHGEGGEGAPTFPSVNDT 584

## RESULT 7

Q8LSN2 PRELIMINARY; PRT; 579 AA.  
AC Q8LSN2;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE AP2/EREBP transcription factor BABY BOOM2.  
GN BBW2.  
OS Brassica napus (Rape).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Topas;  
RA Boutilier K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,  
RA Zhang L., Hattori J., Ouellet T., Liu C.-M., Miki B.L.A.,  
RA Custers J.B.M., van Lookeren Campagne M.M.;  
RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a  
RT conversion from vegetative to embryonic growth."  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
EMBL: AF317905; AAM33801.1;  
DR GO:0005634; C:nucleus; IEA.  
DR GO:0003700; F:transcription factor activity; IEA.  
DR GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro: IPR001471; TF\_ERF.  
DR Pfam: PF00847; AP2-domain; 2.  
DR PRINTS: PR00367; ETHRSPELEMT.  
DR PRODOM: PD001423; TF\_ERF; 2.  
DR SMART: SM00380; AP2; 2.  
SQ SEQUENCE 579 AA; 64263B2B06D3BBA CRC64;

Query Match 27.0%; Score 947; DB 10; Length 579;  
Best Local Similarity 39.1%; Pred. No. 5, 9e-60;  
Matches 243; Conservative 65; Mismatches 155; Indels 158; Gaps 22;

QY 15 HNMVGLSPH-----HMKMEATSAATVPT-----TFMNSQSLSNFGMCYGV 59

```

Db 2 NNNWLGFSLSPEQHHRKDVCSSTTTAVDAGECYCDPTAASDESSAIGTSPSPGV 61
QY 60 -----GENGNHSPPLTWPLKSDGSLCILEALKRSQVQVWVPTSSPKLEDELGATMGT 114
Db 62 VLDATFRDNNSHS-----RDWDINGSAC--NNIHND-----ODGPKLENFLGRTT----- 105
QY 115 EYGSHERGLSDSIYNSQNAEAQPNRDLISQPFROQGHMSVQTHPYSGLAGHLYQAP 174
Db 106 -----TINNTENV-----GDI-----DGGCYG----- 124
QY 175 LBEETKETHVSDCSLMPQWTEGLKXWVAPTRFEFTHQVLEQGMNCGMNERNGVSLG 224
Db 125 -----GDDGGGSLGLSM--IKTW-----LRNQPDVNDVNOENG--NG 158
QY 235 SVGGCELGLSLISMSPGSSSCVTAPSGTDSVAVDAKRGHAKLGQKQPVHRKSIDTFGQ 234
Db 159 AKG-----LSLSMNSTSCDNNNSNNLVAQKTIIDS-----VEATPKTIESFGQ 206
QY 235 RTSQYRGVTRHRTGRYEAHLMDNSCKEGQTRKGRQVYLGQYMEKAAAYDLAALXY 354
Db 207 RTSIYRGVTRHRTGRYEAHLMDNSCKEGQTRKGRQVYLGQYMEKAAAYDLAALXY 266
QY 355 WGPSTHINSIENYQVLEEMKMSRQRYAHLRKKSGFSRGASITRGVTRHHQGRWQ 414
Db 267 WGTITTFPMSEYEKEIEEMKMTROBYVASLRKSSGFSRGASITRGVTRHHQGRWQ 326
QY 415 ARIGVANKDLYLGTFSTOEBAEAAYDVAIKFRGANAVTNFDSRYDVERIMASNNL 474
Db 327 ARIGVANKDLYLGTFSTOEBAEAAYDVAIKFRGLAVTNFDMNRVYVKAIIIESPLP 366
QY 475 AGEIARRKNDPRKNDIDYNSVYTSVNMETVQVQAGNNNDSEMKVLEFNHS-- 532
Db 387 IGSAKRRIKE--ARRPVSMWTISNNVSESE-----NNASQMAAYQHQGV 432
QY 533 -----QOOQANGSGDOKIMNCGNYRN--SAFMAIQLDILIGISVSGQHMLDESSKIGT 586
Db 433 DLSLLQHQERYNG--YYNGNLSSESARACFQK-----DDQHFL----- 473
QY 587 HFSNTSLVTSLSSEASPE 607
Db 474 --SNTQSLMTNIDHQSVSDD 492

RESULT 8
ID 08L3U3 PRELIMINARY; PRT; 579 AA.
AC 08L3U3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE AP2/ERBP transcription factor BABY BOOM.
GN BEM1
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Brassica.
OX NCBI_Taxid=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Topsa;
RA Boutiller T., Offringa R., Sharma V.K., Kieft H., van Lammereen A.A.M.,
RA Cuellet J., Zhang L., Hattori J., Liu C.-M., Miki B.L.A.,
RA Custers J.B.M., van Lookeren Campagne M.M.;
RT "Ecotypic expression of the Brassica napus BABY BOOM gene triggers a
RT conversion from vegetative to embryonic growth."
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317904; AAM33800.1;
DR EMBL; AF317906; AAM33802.1;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.

```

```

DR PRINTS; PR00367; ETRHSPLENT.
DR Prodom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2_2.
DR SEQUENCE 579 AA; 63962 MW; EA2AF4734500D72 CRC64;

Query Match 26.9%; Score 942; DB 10; Length 579;
Best Local Similarity 38.9%; Pred. No. 1,4e-59;
Matches 243; Conservative 61; Mismatches 156; Indels 164; Gaps 21;

QY 15 NNNWLGFSLSPEQHHRKDVCSSTTTAVDAGECYCDPTAASDESSAIGTSPSPGV 59
Db 2 NNNWLGFSLSPEQHHRKDVCSSTTTAVDAGECYCDPTAASDESSAIGTSPSPGV 61
QY 60 -----GENGNHSPPLTWPLKSDGSLCILEALKRSQVQVWVPTSSPKLEDELGATMGT 114
Db 62 VLDATFRDNNSHS-----RDWDINGSACNNIHND-----ODGPKLENFLGRTT----- 105
QY 115 EYGSHERGLSDSIYNSQNAEAQPNRDLISQPFROQGHMSVQTHPYSGLAGHLYQAP 174
Db 106 -----TINNTENV-----GDI-----DGGCYG----- 124
QY 175 LBEETKETHVSDCSLMPQWTEGLKXWVAPTRFEFTHQVLEQGMNCGMNERNGVSLG 234
Db 125 -----GDDGGGSLGLSM--IKTW-----LRNQPDVNDVNOENG-- 156
QY 235 SVGGCELGLSLISMSPGSSSCVTAPSGTDSVAVDAKRGHAKLGQKQPVHRKSIDTFGQ 294
Db 157 ---NAAKLSLSNM--SSTSCDNNNSNNNVAQKTIIDS---VEATPKTIESFGQ 206
QY 295 RTSQYRGVTRHRTGRYEAHLMDNSCKEGQTRKGRQVYLGQYMEKAAAYDLAALXY 354
Db 207 RTSIYRGVTRHRTGRYEAHLMDNSCKEGQTRKGRQVYLGQYMEKAAAYDLAALXY 266
QY 355 WGPSTHINSIENYQVLEEMKMSRQRYAHLRKKSGFSRGASITRGVTRHHQGRWQ 414
Db 267 WGTITTFPMSEYEKEIEEMKMTROBYVASLRKSSGFSRGASITRGVTRHHQGRWQ 326
QY 415 ARIGVANKDLYLGTFSTOEBAEAAYDVAIKFRGANAVTNFDSRYDVERIMASNNL 474
Db 327 ARIGVANKDLYLGTFSTOEBAEAAYDVAIKFRGLAVTNFDMNRVYVKAIIIESPLP 366
QY 475 AGEIARRKNDPRKNDIDYNSVYTSVNMETVQVQAGNNNDSEMKVLEFNHS-- 531
Db 387 IGSAKRRIKE--ARRPVSMWTISNNVSESESENSAGQNAAYQHQ 429
QY 532 S-----QOOQANGSGDOKIMNCGNYRN--SAFMAIQLDILIGISVSGQHMLDESSK 583
Db 430 QGVDSLHGHQERYNG--YYNGNLSSESARACFQK-----DDQHFL----- 473
QY 584 IGFHSNTSLVTSLSSEASPE 607
Db 474 --SNTQSLMTNIDHQSVSDD 492

RESULT 9
ID 09CAH3 PRELIMINARY; PRT; 425 AA.
AC 09CAH3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Putative AP2 domain transcription factor.
GN F88P2.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=2106719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.T., Federspiel N.A., Kaul S.,

```

RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao O., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Elgu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurta D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 RA Miltschev J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,  
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utecherback T., Van Aken S., Vayberg M., Vysotskaya V.S., Walker M.,  
 RA "Yu G., Fraser C.M., Venter J.C., Davis R.W.,  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.", 816-820(2000).  
 RL Nature 408:816-820(2000).  
 DR EMBL: AC010926; AAC51860.1; -.  
 DR PIR: B96750; B96750.  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003700; F:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRSPLELMT.  
 DR ProDom: PD001423; TF\_ERF; 2.  
 DR SMART: SM00380; AP2; 2.  
 SQ SEQUENCE 425 AA; 48205 MW; C191BA52D6C5ACBB CRC64;

Query Match 26.4%; Score 923.5; DB 10; Length 425;  
 Best Local Similarity 43.4%; Pred. No. 1.9e-58;  
 Matches 214; Conservative 60; Mismatches 108; Indels 111; Gaps 16;

QY 18 WLGLSLSPHMKM-----EATSAATVPTTFYMSPSGSHLNFQMGYGVGENGNFSLTV 71  
 DB 4 WLGLSLTPPRICNSEEBELRHDSGLTF-----THRPVWR 42  
 QY 72 MPLKDGSLICILEAKRSOTQVWPTSSPKLEDFLGATGTGTHEYSGHERGLSDIYNN 131  
 DB 43 YDINPDH-----HHHEDV-----PKVEDLLS-----NSHQTYPIN--HN 76  
 QY 132 SQNAEAQPNRDLISQPRQOGHMSVQTHPYYSGLACHGLYQAPLEETKETHVSDCSL 191  
 DB 77 QTNVACTTVYVRLNP-----GYL-----LHDQVVTVPYH--PNT 109  
 QY 192 MPQMTG-----LKNWV--APTRFSTHQVLEQOMNCGM--GNERNVSL 233  
 DB 110 DENLSNDYGFPERVGSVAFKSMLEQGPAPPLSHVYTEAGISNNISHPSNBETGYNT 169  
 QY 234 -GSVGGELQSLSLSPSGSSCVTAPAGT---DSVAVDPAK-KRGHAKLGOKOPVARK 287  
 DB 170 NSGM-----LSLASHGACSDLINESVSAKVEEPKVDKRLVYVKKQVYESVPRK 222  
 QY 288 SIDTFQORTSQYRGVTRRWTRGRYEAHLWDSCKEGQTRKRGQVYLLGQYDMEKAAAY 347  
 DB 223 SVDSYQORTSQYRGVTRRWTRGRYEAHLWDSCKEGQTRKRGQVYLLGQYDMEKAAAY 282  
 QY 348 DLAAALYKWPSTHINSIENYQVLEEMKMSROBYVHLRKSGSGFRGASIVYGVTRH 407  
 DB 283 DLAAALYKWPSTHINSIENYQVLEEMKMSROBYVHLRKSGSGFRGASIVYGVTRH 342  
 QY 408 HOGHGMQARIGRYAGNKDYLIGTFTQEEBAAYVAALIKFRGANAVTNPDISRYVERI 467  
 DB 343 HOGHGMQARIGRYAGNKDYLIGTFTQEEBAAYVAALIKFRGANAVTNPDINRYVERI 402  
 QY 468 MASSNLAGEIAR 480  
 DB 403 CSSSTIVDSDAQ 415

ID Q8LGQ3 PRELIMINARY; PRT; 597 AA.  
 AC Q8LGQ3;  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 GN Ovule development aintegumenta-like protein BNM3.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_Taxid=4530;  
 RA [1]  
 RA SEQUENCE FROM N.A.  
 RA Bi X.-Z.;  
 RT "Cloning and identification of two ovule development proteins,  
 RT aintegumenta-like protein in rice (Oryza sativa)."  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY062180; AAL47205.1; -.  
 DR Gramene: Q8LGQ3; -.  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003700; F:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRSPLELMT.  
 DR ProDom: PD001423; TF\_ERF; 2.  
 DR SMART: SM00380; AP2; 2.  
 SQ SEQUENCE 597 AA; 62198 MW; F856BEC99ADE25B CRC64;

Query Match 26.3%; Score 919.5; DB 10; Length 597;  
 Best Local Similarity 36.0%; Pred. No. 6.1e-58;  
 Matches 253; Conservative 79; Mismatches 206; Indels 165; Gaps 24;

QY 16 HNWLGSLSPHMKM-----EATSAATVPTTFYMSPSGSHLNFQMGYGVGENGNFSLTV 75  
 DB 5 HNWLGSLTPPRICNSEEBELRHDSGLTF-----THRPVWR 42  
 QY 76 SD---SLCILEALKRSOTQVWPTSSPKLEDFLGATGTGTHEYSGHERGLSDIYNN 132  
 DB 38 GDSSTDVCF-----NIPQAH-----STPAIGNGGIGLSMIKWLAS 75  
 QY 133 QNAEAQPNRDLISQPRQOGHMSVQTHPYYSGLACHGL--YQAPLEETKETHVSDCSL 190  
 DB 76 QPA-PQP-AQALSLSMNMAAGTTTAAQGGAAALLAGERGRTTPASESLSTAHHGTTAT 133  
 QY 191 LMPQMTGELKNWVAPTRFSTHQVLEQOMNCGMNERGVSLGSGCELQSLSLMSRP 250  
 DB 134 ---MAGRKET-----NEBSSGSAAY-----VAV 155  
 QY 251 GSGSSCVTAPAGTDSVAVDPAKRGHAKLGOKOPVARKSIDTFQORTSQYRGVTRRWTR 310  
 DB 156 GSE-----SGSGSAVVEA-----GAAAAAARKSVDTFGQRTSIRGVTRRWTR 200  
 QY 311 YEHLWDSCKEGQTRKRGQVYLLGQYDMEKAAAYDLAAALYKWPSTHINSIENYQV 370  
 DB 201 YEHLWDSCKEGQTRKRGQVYLLGQYDMEKAAAYDLAAALYKWPSTHINSIENYQV 257  
 QY 371 QLEEMKMSROBYVHLRKSGSGFRGASIVYGVTRRHOGHGMQARIGRYAGNKDYLIGT 430  
 DB 258 ELEMKMTQRFVVASIRKSGSGFRGASIVYGVTRRHOGHGMQARIGRYAGNKDYLIGT 317  
 QY 431 FSTQEEBAAYVAALIKFRGANAVTNPDISRYVERIYVSSNLAGEIARRKKNQPNRKN 490  
 DB 318 FSTQEEBAAYVAALIKFRGANAVTNPDISRYVERIYVSSNLAGEIARRKKNQPNRKN 377  
 QY 491 -DIDYNSVVTSVNNEETVQVQAGNNNNENDESMKNTLVN-----HP----- 531  
 DB 378 YDVG---RASHGAGGCAVAAYGHHSAAAMPPIAFOAAAPPPHAGLVHPYAQPL 434  
 QY 532 ---SQOQO-----ANGNSDQKIMNG--NYNSAFSMALODLIGDSV-----GSG 573  
 DB 435 RGMCKEODHAAVIAAASLQDLHLHLNLGAAAAAHDFFSQMQOQHGLSDINASLEHSTG 494

QY 574 QHNL---DESSKIGTHFNSISLVYSLSSREASPEKGPILLFPMETKIV--NPI 628  
 Db 495 SNSVYNGDNGGGGGYIMAPMSAVSATATAVASHHGGGGKQVOMQYDYLVDADAY 554  
 QY 629 GTSVTSMLEPTVQMRPSPALISLH-----LPVFASWTDT 663  
 Db 555 GGGGAGRMPSWAMTPASAPATISSDMTGYCHGAQLFSVWNTQ 597

RESULT 11  
 QY6D4 PRELIMINARY; PRT; 655 AA.  
 ID QY6D4; 07X6D4;  
 AC 07X6D4;  
 DT 01-OCT-2003 (Tremblrel. 25, Created)  
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
 DE OSUNBA0014K14.16 protein (OSUNB0116K07.8 protein).  
 GN OSUNBA0014K14.16 OR OSUNB0116K07.8.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbiaceae; Oryzae; Oryza.  
 CX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,  
 Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,  
 Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,  
 Shao H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,  
 Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,  
 Chen L., Fan D.L., Meng Q.J., Zhang L., Lu Y.Q., Yu S.Q., Liu X.H.,  
 Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,  
 Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao F., Zhang L., Wu M.,  
 Zhang R.Q., Guan J.F., Hong G.F.,  
 RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
 Lu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Meng Q.J., Zhang L.,  
 Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,  
 Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen X.Y.,  
 Hao F., Zhang L., Wu M., Zhang R.Q., Guan J.F., Tu Y.F., Jia J.,  
 Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
 Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
 Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
 Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.,  
 RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL606604; CAB02944.1;  
 DR EMBL; AL731638; CAB05555.1;  
 SQ SEQUENCE 655 AA; 67291 MW; D6FCDD0F729A6933 CRC64;

Query Match 26.2%; Score 918.5; DB 10; Length 655;  
 Best Local Similarity 42.8%; Pred. No. 8.2e-58;  
 Matches 227; Conservative 51; Mismatches 117; Indels 135; Gaps 18;

QY 17 NMLGFSLS-----PHMKMTASATVPTTFYSPSGSHSNFGMCGVGENGNFSPITLV 71  
 Db 6 NMLGFSLSGCGNPGHONGSPSA-----GDAIDISGSDPFGULTPPAHITGA- 56  
 QY 72 MPLKSDGSLCTLELKR-----SQTQWMPFS----- 98  
 Db 57 ---GEDAPYGYMDAFNGGTHETDPMANGLDYGSGSSDLSMLVSGSGGRRTVADGVGE 113  
 QY 99 SPKLEDFLGATM-----GTHEY-----GSHERG-----LSLDSIYNSON 134  
 Db 114 APRLENFLDNGSPFDVHGAAGVLYGSAVGAGVYNGCGCGGTIELSMIKTLRNSQ 173  
 QY 135 AEAQPNRDLISQPRQGHNSVQTHPYISGLACHGLYQAPLEBETTKETHVDCSLMPQ 194  
 Db 174 SQQPP-----SPF---QAADQGMSTDAASSYAC-----SDVLVSCCG----- 209  
 QY 195 MTEGLKTVAPTRFETHQVLEQNMNCGKNGERNNGVSLGVSQCEGLQSLSTMSPGSQS 254

Db 210 ---GAGGTA-----SSHQGLALMSGT-----SVAAAGG--CAVVAABSSSEN 251  
 QY 255 SCVTAPSGTDSVAVDKKRGAHAKQOKOPVARKSIDTFOQRTSOYRGVTRRWGRYBAH 314  
 Db 252 KRVDSFGG---AVDG-----AVPKSIDTFOQRTSIYRGVTRRWGRYBAH 295  
 QY 315 LMDNSCKKSGQTRKSGQVYLGVDMEKAPRAYDLAAIKYMGSPRHINFSIENVOVLEE 374  
 Db 296 LMDNSCKRSGQSKGRQ---GGYDKEDKAPRAYDLAAIKYMGCTTTTTFPMSNEKLEEE 352  
 QY 375 MKNMSRQEVYAHLRKSGSPSGASITRGVTRHGHGWQARIGFVAGNKDLYLGTSQ 434  
 Db 353 MKMTROEVYIAHLRNSGSGFSGASKYRGVTRHGHGWQARIGFVAGNKDLYLGTSQ 412  
 QY 435 EEAABAYDAVATKFGANAVTNPDISRYDVERIMSSNLLAGELARKKD 484  
 Db 413 EEAABAYDAVATKFGANAVTNPDISRYDVERIMSSNLLAGELARKKD 462

RESULT 12  
 QY6F42 PRELIMINARY; PRT; 581 AA.  
 ID QY6F42; 09LF42;  
 AC 09LF42;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DE 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Ovine development protein atintegumenta-like protein.  
 GN 110B6.90.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eumossids II; Brassicales; Brassicaceae; Arabidopsis.  
 CX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,  
 Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,  
 RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RA EUR Arabidopsis sequencing project;  
 RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL91142; CAC01738.1;  
 RL PIR; T51580; T51580.  
 DR GO; GO:0005634; Cytosol; IEA.  
 DR GO; GO:0003700; P; transcription factor activity; IEA.  
 DR GO; GO:0006355; P; regulation of transcription, DNA-dependent; IEA.  
 DR Interpro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS; PRO0367; ETHRSPELEMT.  
 DR PRODOM; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 581 AA; 64170 MW; 292A8ACC29BB79D CRC64;

Query Match 35.2%; Score 916.5; DB 10; Length 581;  
 Best Local Similarity 26.2%; Pred. No. 9.6e-58;  
 Matches 255; Conservative 76; Mismatches 170; Indels 223; Gaps 25;

QY 16 NMLGFSLSPPHMKMTASATVPTTFYSPSGSHSNFGMCGVGENGNFSPITLVPLK 75  
 Db 5 NMLGFSLSPPH-----DONH-----HRTVDSTTRTAVD 34  
 QY 76 SDGSLCI-LEALKRSQTVMPPTSP---KLEDFLGATMGTHEYSHER-----GLSLD 126  
 Db 35 VAGGYCFDLAASDSSSAVQTSFSPFGVTLBAF-----TRNNNSGRWMDNGACN 87  
 QY 127 SIYYSQNAEAPNDLISQPRQGHNSVQTHPYISGLACHGLYQAPLEBETTKETHVS 186  
 Db 88 NINNEQON----- 96  
 QY 187 DCSLMPQMTBGLKTVAPTRFETHQVLEQNMNCGKNGERNNGVSLG----- 234

Db 97 -----BK-----LENFGRITTTIYNTNETVVDGNGDGGGGGGLGSLIKWTLSNH 146  
 Qy 235 SVGGGELQ-----SLSLMSPGSQSCVTAPSGTDSVAVDAKKRGHAKLQKQPVHRKSI 289  
 Db 147 SVAAHNGQDNGNGARGSLSLMSNSTSDSNNNNDV---VOEKITVDVETTP--KXTI 201  
 Qy 290 DTEGRTSQRTGRTYRHAHTYRHAHLMDNSCKREGQTRGRQYVLLGYTMEKARAYDL 349  
 Db 202 ESFGKRTSTYRGTYRHAHTYRHAHLMDNSCKREGQTRGRQ---GGYDKKARAYDL 258  
 Qy 350 AALKYMGSTHINFSIENYQVLEEMKMSRQCYVAHLRKSGSFGSGASIRGYTRHQ 409  
 Db 259 AALKYMGSTHINFSIENYQVLEEMKMSRQCYVAHLRKSGSFGSGASIRGYTRHQ 318  
 Qy 410 HGRQARIGYVAQKDIYLGTFSTQEEAAAYVAALKRGANATNPDISTYDERIMA 469  
 Db 319 HGRQARIGYVAQKDIYLGTFSTQEEAAAYVAALKRGASAVTNPDNMYKALLE 378  
 Qy 470 SSNLAGEARRRKKDNPKNKDIDYKSVYTSVNNETVQVQAGN--NNNENDEMKAVL 527  
 Db 379 SPSP.LIGSSAKRLKD-----VNNVPMAMT.SNNVSEBANNVSGMONTA 421  
 Qy 528 FPHPS-----QQQANGSGSQKIMNGCNTYNSAFSALODLIGDSVSGQNMIDE 580  
 Db 422 FOHHQMDLSLQQQDERRVY---YNGNGLSTESTRVCFK-----QEEBQCHFLNS 471  
 Qy 581 SSKIGT--HFSNTSLVTSLSRSREASPEKRGPSLLFPMPMETKIVNPIGTSVTLPS 638  
 Db 472 PSHMTNVDDHSSSDSVTCGVVSYGGYGPAI-----PVGTSV-NYDPF 517  
 Qy 639 PTV-----QMRSP-----AISLSH-----LPVPS 659  
 Db 518 TAAEIAVNAHNYTAAHQOQOQOQOQOQSGPDPEVALSNHNNMTFHGGGEGAGAPTSV 577  
 Qy 660 WTD 663  
 Db 578 WNDT 581

RESULT 13  
 Q9F1E2 PRELIMINARY; PRT; 555 AA.  
 AC Q9F1E2; (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Similarity to unknown protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RX MEDLINE=9156233; Pubmed=10048468;  
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kobani H., Miyajima N.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.  
 RT Sequence features of the regions of 1,081,956 bp covered by seventeen  
 RT physically assigned pl and YAC clones.";  
 RL DNA Res 5:379-391(1998).  
 DR EMBL, AB016891, BAB08476.1;  
 DR GO:GO:0005654; C:nucleus; IEA.  
 DR GO:GO:0003700; F:transcription factor activity; IEA.  
 DR GO:GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001471; TF\_SRF.  
 DR Pfam: PF00847; AP2-domain; 2.  
 DR PRINTS: PR00367; ETRHSPLEMT.  
 DR PRODOM: PD001423; TF\_SRF; 2.  
 DR SMART: SM00380; AP2\_2.  
 QO SEQUENCE 555 AA; 59930 MW; 881FL43370E3A3A1 CRC64;

Query Match 25.7%; Score 900.5; DB 10; Length 555;  
 Best Local Similarity 36.0%; Pred.No.1,3e-56;  
 Matches 245; Conservative 78; Mismatches 182; Indels 175; Gaps 22;  
 Qy 7 SNNTDGNNHNNLGFSL---PHMKGEATSAATVTTTMYSPSQHLENFQKCVGANG 63  
 Db 19 SPSSSSSSQNNLSTSLNNNNNNFSSNPNTLSTSDHHHPHSLF-----68  
 Qy 64 NFHSPLTPML-KSDGLCILALRSQTYV-VPTSSPKEDFLGATMGTHGSHR 121  
 Db 69 ---QAFSTSPVRQDSPFVSPS---DATAVLSTVPCGPKLENFPGCASST-----114  
 Qy 122 GLSLDIYNNQNAEQPRDLISQPFQGGHMSYQTHPYSGLAChLYQAPLEETTK 181  
 Db 115 -----TTPMQ-----VQS---LGGVFSDDLPPHPSAA 144  
 Qy 182 ETHVSDCSLMPQMTGLKNVVAPTREFSTHQVLEQMGNGMERNVSLGSGCEL 241  
 Db 145 EIVDSLKSLIASF-----LGVSGHSEVSVAHQ 176  
 Qy 242 QSLSLMSPGSQSCVTAPSGTDSVAVDAKKRGHAKLQKQPVHRKSIDTEGRTSQYRG 301  
 Db 177 QRPPLVSEAS-----PTEKQVESFGQRTISTYRG 206  
 Qy 302 VTRHRTGYRHAHLMDNSCKREGQTRGRQYVLLGYTMEKARAYDLAALKYMGPSYTHI 361  
 Db 207 VTRHRTGYRHAHLMDNSCKREGQTRGRQ---GGYDKKARAYDLAALKYMGPTTTT 263  
 Qy 362 NFSIEYQVLEEMKMSRQCYVAHLRKSGSFGSGASIRGYTRHQHGRQARIGVA 421  
 Db 264 NFPISTYSELEEMKMTROEFVSLRKSSGFSGASMTGYTRHGHGQWQARIGVA 323  
 Qy 422 GNKDIYLGTFSTQEEAAAYVAALKRGANATNPDISTYDERIMASNL-LAGEIAR 480  
 Db 324 GNKDIYLGTFSTQEEAAAYVAALKRGANATNPDISTYDERIMASNL-LAGEIAR 382  
 Qy 481 RKDNDPRNKDIDYKNS-----VYTSVNNETVQVQAG--NNNENDEMKAVL 525  
 Db 383 PSPATADKTYDLSBSPSLTTPSLTFNATPVNDGGTFYHTGIDIKDPADHYWSN 442  
 Qy 526 VLFNHSQOQANGSGSQKIMNGCNTYNSAFSALODL-----GIDSVSGQNM 577  
 Db 443 IF-----GFQANPRAEMRPLANFGSDLHNPSPGYAIVPQOENNGGFSVSDGYN 496  
 Qy 578 LBESSKIGHFNTSLVTSLSRSREASPEKRGPSLLFPMPMETKIVNPIGTSVTLPS 637  
 Db 497 HSAAS-----NPVSAIPUSSTTTWNGNGEGYGN-----INWLNINNIS--535  
 Qy 638 SPTVQMRSPPAISLHPVF 657  
 Db 536 --SYCTAKS-NISVLTPTVF 552

RESULT 14  
 Q9L137 PRELIMINARY; PRT; 540 AA.  
 AC Q9L137;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Similarity to unknown protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Columbia;  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 QO SEQUENCE FROM N.A.

RC STRAIN=Columbia;  
 RX MEDLINE=20277480; PubMed=10819329;  
 RA Nakamura Y;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
 RT clones";  
 RL DNA Res. 7:131-135(2000).  
 DR EMBL; AB025629; BAB02492.1; -;  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR01471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR PRINTS; PR00367; ETRKSPLEMT.  
 DR ProDom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 540 AA; 59066 MW; 0A013B0735176DDD CRC64;

Query Match 25.5%; Score 893.5; DB 10; Length 540;  
 Best Local Similarity 38.2%; Pred. No. 4e-56;  
 Matches 238; Conservative 74; Mismatches 142; Indels 169; Gaps 23;

QY 100 PKLEDFLGATGCTHGYGHERGLSDSTYNSQNAQPNEDLSQPFROGHMSVQTH 159  
 DB 27 PKVADPLG--VSKPDENQSNHLVAYVDSDFPHNS-LMP-----SYQSN 68  
 QY 160 PYYSLAGHGLVQAPLEETTKETHVSDSSSLMPQMTGLKQWVAFTSESTHQVLEQQ 219  
 DB 69 -----DVVVAACDSNTFNK-----SSYHLOQSA 92  
 QY 220 MNCQGNERNVSLGSGGELQSLSNSSPGSSCVTAPSGTDSVAVDKRRHAYLG 279  
 DB 93 HN-----LQSLTSMGTTAGNNVVDKAPSETTGDNAGCALAYVE 133  
 QY 280 QKQPVHRKSIDTGCQTSQYRGVTRHRTGRYEAHLMNSCKKEQTRKGRQVYLGDM 339  
 DB 134 TATP--RRALDTFGQRTSYRGVTRHRTGRYEAHLMNSCKKEQTRKGRQVYLGDM 188  
 QY 340 EERKAPAYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEVYAHLRKSGSFRGAS 399  
 DB 189 EDKASSYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEVYAHLRKSGSFRGAS 248  
 QY 400 IYGVTRHOGHGMQARIGRVAGNKDLYIGTSTQEEAAEAVDAIKRGANATNDI 459  
 DB 249 MYRGVTRHOGHGMQARIGRVAGNKDLYIGTSTQEEAAEAVDAIKRGANATNDI 308  
 QY 460 SRVDVERIMASSNL-LAGE-LARKCD-----NDRNKDID----- 493  
 DB 309 NRVDVAILLESSTLPIGGGAARLKEAQLLESERKEAEMIALGSGFYGGSGSTGSGST 368  
 QY 494 -----YKSY-----VTSVNNBETVQVQANNNNENDEMKVLPNNHSSQ----- 533  
 DB 369 SSRLOQOPYPYLSIQOPEPLSLQNNND--ISHYNNNNNAHDS-----SENHHSYTOTQAH 421  
 QY 534 -----OOQANGNSDQKIMNCGYNSAFSMALODLIGDSVGSQAHMLDESSXI 584  
 DB 422 LHQCTNNYLOOQSSQ--SQQLTNAIYHSNPAL--LHGIVSTSYIDNNNN-----GSSS 472  
 QY 585 GTHESNTSLVTSI-----SSSRPASEPKRGPSLLPMPMETKIYNPIGTSTVSWLSPF 639  
 DB 473 GSY--NTAAFLGNHGIGIGSSSTVGSTEE-----PPTVTDYDMSDGTGIGYSGMTSE 524  
 QY 640 TYQMRPSPALSLSLPVEASMTD 662  
 DB 525 SVQ-GSNPG-----GVFTYMN 540

RESULT 15  
 Q41832 PRELIMINARY; PRT; 485 AA.  
 AC Q41832  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Of protein.  
 GN ORF.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.  
 OC NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Black Mexican Sweet, TISSUE=Endosperm,  
 RA Daniel T.J., Fordham-Skelton A.P., Vergani P., Edwards R.;  
 RT "Isolation of a Maize cDNA (Accession No. Z47554) Encoding APETALA2-  
 RT Like Binding Domains by Complementation Cloning of an L-Isoaspartyl  
 RT Methyltransferase-Deficient Mutant of Escherichia coli (PGR 96-013).";  
 RL Plant Physiol. 110:1435-1435(1996).  
 DR EMBL; Z47554; CAA87634.1; -;  
 DR PIR; T03638; T03638.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR ProDom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 485 AA; 52391 MW; 2B165E3D17FC8B3A CRC64;

Query Match 25.4%; Score 890; DB 10; Length 485;  
 Best Local Similarity 44.2%; Pred. No. 6.1e-56;  
 Matches 200; Conservative 45; Mismatches 94; Indels 114; Gaps 10;

QY 187 DCSLMPQMTGLKQWVAFTSESTHQVLEQQMNC--MGRKRGV-----LG 234  
 DB 4 DMSAYPH-----HMLSFSLNNYHGLLEAFSSNGTFLGDQGAVERSPFTVEDFLG 57  
 QY 235 SVG-----CGELQSLSNSSPGSS-----CVTAPSGT 263  
 DB 58 GVGCVGAPRRRLQIRITSCASCSITRFLRHPAQSCTVGEPLSRFLAMMSST 117  
 QY 264 DSYAVDAKKRGAHKLQKQPVHRKSIDTGCQTSQYRGVTRHRTGRYEAHLMNSCKE 323  
 DB 118 DVAAEASDQAS-----RSAETFGQRTSYRGVTRHRTGRYEAHLMNSCKE 165  
 QY 324 GQTFKRGVYLGVDMEKARAYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEV 383  
 DB 166 GQSRKRGVYLGVDMEKARAYDLAALKYGPSTHINFSIENTYQVLEEMKMSRQEV 225  
 QY 384 VAHLRRKSGSFRGASIRGVTRHOGHGMQARIGRVAGNKDLYIGTSTQEEAAEAYDV 443  
 DB 226 IASLRKSGSFRGASIRGVTRHOGHGMQARIGRVAGNKDLYIGTSTQEEAAEAYDI 285  
 QY 444 AAIKFRGANVTNDDISRYVERIM-----ASSNLAG 476  
 DB 286 AAIKFRGANVTNDDISRYVERIM-----ASSNLAG 345  
 QY 477 ELARKKNDPNRNDI-----DYNSVVTYNNBETVQVQANN 515  
 DB 346 AAASQATMPSEKDYMSLLALHYQOOEORQPAQAYAYAGSGGVNVDFTMTGSSGSN 405  
 QY 516 NNENDS-EW-----KMTLFNHPSCQOQANGGSD 543  
 DB 406 NNTGSGVWGATSGAVVGOQSSSKQNGYASN 438

Search completed: March 9, 2004, 10:47:03  
 Job time : 49 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:29:58 ; Search time 60 Seconds  
(without alignments)  
3122.150 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502  
Sequence: 1 MKRINESNNITDDGNNHWLG.....RPSPAISLSHPFASMTDR 663

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: geneeqp1980s:\*
- 2: geneeqp1990s:\*
- 3: geneeqp2000s:\*
- 4: geneeqp2001s:\*
- 5: geneeqp2002s:\*
- 6: geneeqp2003as:\*
- 7: geneeqp2003bs:\*
- 8: geneeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3502	100.0	663	5	ABR79636 Soybean A
2	1715	49.0	665	5	ABR79637 Soybean A
3	1408.5	40.2	642	5	ABR79638 Rice AINT
4	1349.5	38.5	669	5	ABR79639 Rice AINT
5	1338	37.9	555	3	ABR07724 An Arabid
6	1328	37.9	555	3	AA31380 Arabidops
7	1328	37.9	555	6	ABR40832 Arabidops
8	1328	37.9	555	6	ABR40830 Arabidops
9	1307	37.3	548	3	AA31382 Arabidops
10	1295.5	37.0	528	3	AA31381 Arabidops
11	1295.5	37.0	529	3	AA31381 Arabidops
12	1180.5	33.7	585	5	ABR79640 Cotton Ai
13	1061.5	30.3	557	6	ABR40856 Glycine m
14	1059.5	29.9	710	6	ABR40768 Glycine m
15	1046.5	29.2	384	6	ABR40768 Glycine m
16	1021.5	27.3	392	6	ABR40770 Glycine m
17	956	27.1	510	6	ABR40877 Glycine m
18	948.5	27.1	510	6	ABR40877 Glycine m
19	948	27.1	574	7	ADD30879 Plant yie
20	948	27.1	574	7	ADD30879 Plant yie
21	948	26.9	574	7	ADD30834 Plant yie
22	942	26.9	579	4	ABR50873 Brassica
23	939.5	26.8	415	7	ADD30560 Plant yie
24	936	26.7	553	6	ABR40771 Glycine m

25	934	26.7	538	6	ABR40854 Goseypium
27	931.5	26.6	470	6	ABR40853 Catalpa s
28	927	26.5	486	6	ABR40767 Glycine m
29	927	26.5	558	7	ADD30881 Plant yie
30	923.5	26.4	425	6	ABR40833 Arabidops
31	922.5	26.3	489	6	ABR40756 Zea mays
32	922.5	26.3	512	6	ABR40852 Glycine m
33	920.5	26.3	530	6	ABR40769 Glycine m
34	918	26.2	469	6	ABR40759 Zea mays
35	911.5	26.0	484	6	ABR40757 Zea mays
36	893.5	25.5	540	6	ABR40837 Arabidops
37	891.5	25.5	498	7	ADD30350 Plant yie
38	890	25.4	485	6	ABR40829 Zea mays
39	890	25.4	485	6	ABR40834 Zea mays
40	879	25.1	516	6	ABR40836 Arabidops
41	874	25.0	566	6	ABR40855 Zea mays
42	831	23.7	592	6	ABR40876 Arabidops
43	717	20.5	160	6	ABR40761 Glycine m
44	696.5	19.9	312	6	ABR40762 Glycine m
45	687	19.6	412	6	ABR40851 Glycine m

## ALIGNMENTS

RESULT 1  
ABR79636  
ABR79636 standard; protein; 663 AA.

XX ABR79636;

XX 21-OCT-2002 (first entry)

XX Soybean AINTEGUMENTA-like polypeptide GmANT1.

XX AINTEGUMENTA; ANT-like polypeptide; GmANT1; soybean; transgenic plant;

XX plant.

XX Glycine max.

XX WO200259332-A2.

XX 01-AUG-2002.

XX 19-DEC-2001; 2001MO-US049294.

XX 21-DEC-2000; 2000US-0257896P.

XX (MONS ) MONSANTO TECHNOLOGY LLC.

XX He SS, Doteon SB;

XX WPI; 2002-599798/64.

XX N-PSDB; ABR4480.

XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in improving agronomic, horticultural, and quality traits of plants, such as increased size of plant organs.

XX Claim 3; Page 130-133; 169pp; English.

The present sequence is the protein sequence of GmANT1, a newly identified AINTEGUMENTA-like (ANT-like) polypeptide of soybean. The sequence is predicted from clone CPR6763 (see ABR4480), which was obtained from a soybean sequence database screening using Arabidopsis ANT as query. Also isolated was clone CPR6762 encoding GmANT2 (see ABR79637). GmANT1 and GmANT2 show homology to ANT in 2 N-terminal AP2 DNA binding domains, but have C-terminal sequences that bear little, if any, homology to ANT although they share conserved segments with each other. The invention provides nucleic acids encoding ANT-like polypeptides comprising, in the N-terminal to C-terminal direction, 2 AP2 DNA binding domain followed by an amino acid subsequence selected from those given in ABR79629-35. ANT-like polypeptides have been identified in soybean.

rice, cotton and corn (see ABB79636-41). Nucleic acids (see ABB84480-86) encoding the ANT-like polypeptides can be used in the construction of transgenic plants, especially corn, soybean, canola, wheat, cotton, tomato or potato (all claimed) having improved agronomic, horticultural or other quality traits, such as increased size of plant organs. These plants are especially useful for production of ethanol or animal feedstuff

Sequence 663 AA;

Query Match 100.0%; Score 3502; DB 5; Length 663;

Best Local Similarity 100.0%; Pred. No. 3, 1e-316; Indels 0; Gaps 0;

Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKRINESNNTDDGNNHNLGFSLSPHMKMEATSAATVPTTTFMSBDSQSHLSNFGMCYGVG 60  
DB 1 MKRINESNNTDDGNNHNLGFSLSPHMKMEATSAATVPTTTFMSBDSQSHLSNFGMCYGVG 60  
QY 61 ENGNTFHSPLTVMPLKSDGSLCTLEALKRSQTVQVWPSSPKLEDFLGATWGTHEYSGHE 120  
DB 61 ENGNTFHSPLTVMPLKSDGSLCTLEALKRSQTVQVWPSSPKLEDFLGATWGTHEYSGHE 120  
QY 121 RGLSLDSIYNSQNAEAQPNRDLISQPRQCGHMSVQTHPYYSGLACHGLYQAPLEBEFTT 180  
DB 121 RGLSLDSIYNSQNAEAQPNRDLISQPRQCGHMSVQTHPYYSGLACHGLYQAPLEBEFTT 180  
QY 181 KETHYSDSCSLMPQMTGKKNVAPTRFSTHQVLEQOMNGMENGVSLSVGCGE 240  
DB 181 KETHYSDSCSLMPQMTGKKNVAPTRFSTHQVLEQOMNGMENGVSLSVGCGE 240  
QY 241 LQSLSLSMSPGSSCCTAPSGTDSVAVDAKRGHAKLQKQPVHRKSIDTFGQRTSQYR 300  
DB 241 LQSLSLSMSPGSSCCTAPSGTDSVAVDAKRGHAKLQKQPVHRKSIDTFGQRTSQYR 300  
QY 301 GYTRHRTGRYEAHLMDNSCKEGRKRGROYLLGYDMEERAAAYLLAALKTWGPSTH 360  
DB 301 GYTRHRTGRYEAHLMDNSCKEGRKRGROYLLGYDMEERAAAYLLAALKTWGPSTH 360  
QY 361 INPSIENYVOLEEMKNMROEYVAHLRKSSGFSRGASITYGVRHHQGRWQARIGRV 420  
DB 361 INPSIENYVOLEEMKNMROEYVAHLRKSSGFSRGASITYGVRHHQGRWQARIGRV 420  
QY 421 AGNKDYLIGTSTOEBAEAAYVAALKEFGANAVTNFDIRYDVERIMASSNLLAGELAR 480  
DB 421 AGNKDYLIGTSTOEBAEAAYVAALKEFGANAVTNFDIRYDVERIMASSNLLAGELAR 480  
QY 481 RKXNDPRKKDIDYNSVTVSVNNBETVOVQAGNNNNENDSEMKNVLFNHPSSQOQOQANGN 540  
DB 481 RKXNDPRKKDIDYNSVTVSVNNBETVOVQAGNNNNENDSEMKNVLFNHPSSQOQOQANGN 540  
QY 541 GSDQKIMNGCNYNASAFNALODLIGIDSVGSQGHMLEDSSKIGTHFENISSLVTSLS 600  
DB 541 GSDQKIMNGCNYNASAFNALODLIGIDSVGSQGHMLEDSSKIGTHFENISSLVTSLS 600  
QY 601 SRASAEKRGPSLLFPMPMETKIYVPIGTSVTSMLPSPTVQMBSPPAISLSHLVPAISM 660  
DB 601 SRASAEKRGPSLLFPMPMETKIYVPIGTSVTSMLPSPTVQMBSPPAISLSHLVPAISM 660  
QY 661 TDT 663  
DB 661 TDT 663

RESULT 2

ABB79637 standard; protein; 665 AA.

XX ABB79637;  
XX 21-OCT-2002 (first entry)  
XX Soybean AINTEGUMENTA-like polypeptide GmANT2.  
XX

KM AINTEGUMENTA; ANT-like polypeptide; GmANT2; soybean; transgenic plant;  
KW plant.

XX Glycine max.

XX WO200259332-A2.

XX 01-AUG-2002.

XX 19-DEC-2001; 2001WO-US049294.

XX 21-DEC-2000; 2000US-0257896P.

XX (MONS ) MONSANTO TECHNOLOGY LLC.

XX He SS, Dotsen SB;

XX WPI; 2002-599798/64.

XX N-PSDB; ABB84481.

XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in improving agronomic, horticultural, and quality traits of plants, such as increased size of plant organs.

XX Claim 3; Page 137-140; 169pp; English.

The present sequence is the protein sequence of GmANT2, a newly identified AINTEGUMENTA-like (ANT-like) polypeptide of soybean. The sequence is predicted from clone CBR6762 (see ABB84481), which was obtained from a soybean sequence database screening using Arabidopsis ANT as query. Also isolated was clone CBR6763 encoding GmANT2 (see ABB79636). GmANT1 and GmANT2 show homology to ANT in 2'-terminal AP2 DNA binding domains, but have C-terminal sequences that bear little, if any, homology to ANT although they share conserved segments with each other. The invention provides nucleic acids encoding ANT-like polypeptides comprising, in the N-terminal direction, 2 AP2 DNA binding domains followed by an amino acid subsequence selected from those given in ABB79629-35. ANT-like polypeptides have been identified in soybean, rice, cotton and corn (see ABB79636-41). Nucleic acids (see ABB84480-86) encoding the ANT-like polypeptides can be used in the construction of transgenic plants, especially corn, soybean, canola, wheat, cotton, tomato or potato (all claimed) having improved agronomic, horticultural or other quality traits, such as increased size of plant organs. These plants are especially useful for production of ethanol or animal feedstuff

SQ Sequence 665 AA;

Query Match 49.0%; Score 1715; DB 5; Length 665;

Best Local Similarity 54.4%; Pred. No. 7, 4e-150;

Matches 379; Conservative 84; Mismatches 166; Indels 68; Gaps 20;

QY 1 MKRINESNNTDDGNNHNLGFSLSPHM-----KMEATSAATVPTTTFMSBDSQSHLSN 52  
DB 1 MKSMENDNDNLDNNQNNWLGFSLSPPMHNIGVSSHSQPSAAAVVPISFHHHTAP--LSS 58  
QY 53 FGMCGYV-GENGNPHSPLTVMPLKSDGSLCTLEALKRSQTVQVWPSSPKLEDFLGATM 111  
DB 53 YGFIYGLAENVGLYSALPIMPLKSDGSLXGLETILRSQQAAMATSTPLNENFLGSEAM 118  
QY 112 GT-HEY--GSHERGSLDSIYNSQNAEAQPNRDLISQPRQ-----QGHMSVQTHPY 162  
DB 112 GTPHHECSATETMPLSLDSVFY-IPSRDRDPNNQTYQNHVQHISTNQOQOQELQAY 177  
QY 163 SGLACHGLYQAPLEBEFTKETHVSDSCSLMPQ-----MTGLKNVAPTRF-STHQ 213  
DB 178 STLNDHMYL-----ESSKQSGTSDNNNLVQNMGGDAPVYGLKSM--EVANFQASHA 230  
QY 214 QVLEQOMNGMGERNGV--SLGVSQGLQSLSLMSPGSSCCTV-----APSGTDSV 266  
DB 231 H--ESKMIIVPHEENAGESGSGISGMAYGDIQSLSLMSPGSSSVTSSHRASPAVVDV 288  
QY 267 AVDAKRGHAKLQKQPVHRKSIDTFGQRTSQYRGVTRHRTGRYEAHLMDNSCKEGRGT 326



```

DB 269 AMDTKKKEKPEKVDKQIVHRKSTIDTFQGRISQYGVTRHMTGTYEHLMDNCKEKGQS 348
QY 327 RKGEVYLIGYDMEKKARAYDALAKYWGSPSTHINSINYQVLEEMQNRQEVVAH 386
DB 349 RKGEVYLIGYDMEKKARAYDALAKYWGSPSTHINPLINYQVLEEMQNRQEVVAH 408
QY 387 LRKSSGSGSRGASLYRGVTRHGHGRQWQARIGRVAGKNDLYLGFSTQEEAAAYDVAI 446
DB 409 LRKSSGSGSRGASLYRGVTRHGHGRQWQARIGRVAGKNDLYLGFSTQEEAAAYDVAI 468
QY 447 KFRGANAVTNFDIRYDVERIMASSNLAGELARRKNDPRNKDIDYKSVTVSV-NNE 505
DB 469 KFRGANAVTNFDIRYDVERIMASSNLAGELARRKNDPRNKDIDYKSVTVSV-NNE 528
QY 506 ETVQVQAGNNNNNDSEMKVLFNHPGQQQOANGSDOKIMNCNGYRNSAFSMALODLI 565
DB 529 EAILMHQKSCSEND-QKRWVLY-QSSQGLEQNPTIE-----SDKTNSFAVALDNMF 580
QY 566 GIDSVGSGQNMDESSKIGTHFNTSGLVTSLSSSSEASPEKRGPSLLFPMPMETKIV 625
DB 581 -----HQVEESSKARTHVSNPSSLATSLSSRSSESPDRSLPMLSGMPSATSKL 631
QY 626 NPIGTSVTSNLPSPVQKRPSPALISLILPYFASMTD 662
DB 632 ATNPNVNVMSDPSPLR-----PALTLPQMPVFAMMTD 664

RESULT 3
ABB79638
ID ABB79638 standard; protein; 642 AA.
XX ABB79638;
AC 21-OCT-2002 (first entry)
XX
DT
XX
DE Rice AINTEGUMENTA-like polypeptide OsANT1.
XX
KM AINTEGUMENTA; ANT-like polypeptide; OsANT1; rice; transgenic plant;
XX plant.
XX
OS Oryza sativa.
XX
PN WO200259332-A2.
XX
PD 01-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US049294.
XX
PR 21-DEC-2000; 2000US-0257896P.
XX
PA (MONS ) MONSANTO TECHNOLOGY LLC.
PI He SS, Dotson SB;
XX
DR MPI: 2002-599798/64.
XX N-PSDB; ABB79638.
XX
PT New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in
PT improving agronomic, horticultural, and quality traits of plants, such as
PT increased size of plant organs.
XX
PS Claim 3; Page 143-146; 16pp; English.
XX
CC The present sequence is the protein sequence of OsANT1, a newly
CC identified AINTEGUMENTA-like (ANT-like) polypeptide of rice. The sequence
CC was predicted from a cDNA clone (see ABB79638) obtained from particle RNA
CC by PCR amplification. OsANT1 cDNA was not amplified from root or leaf
CC RNAs, suggesting tissue-specific expression. Initial OsANT1 clones had
CC been identified in a rice database screening using the newly identified
CC soybean ANT-like polypeptide GmANT1 as query. An additional rice ANT-like
CC polypeptide, OsANT2 (see ABB79639), was also isolated. OsANT1 polypeptide
CC shares high homology with Arabidopsis ANT, GmANT1 and GmANT2 at the N-

```

```

CC terminal AP2 DNA binding domains, shares conserved segments at the N-
CC terminus, and shares conserved segments with GmANT1 and GmANT2, but not
CC with ANT, at the C-terminus. The invention provides nucleic acids
CC encoding ANT-like polypeptides comprising, in the N-terminal to C-
CC terminal direction, 2 AP2 DNA binding domains followed by an amino acid
CC subsequence selected from those given in ABB79629-35. ANT-like
CC polypeptides were identified in soybean, rice, cotton and corn (see
CC ABB79636-41). Nucleic acids (see ABB79636-41) encoding the ANT-like
CC polypeptides can be used in the construction of transgenic plants,
CC especially corn, soybean, canola, wheat, cotton, tomato or potato (all
CC claimed) having improved agronomic, horticultural or other quality
CC traits, such as increased size of plant organs. These plants are
CC especially useful for production of ethanol or animal feedstuff
XX
SQ Sequence 642 AA;

```

```

Query Match 40.2%; Score 1408.5; DB 5; Length 642;
Best Local Similarity 46.5%; Pred. No. 2.4e-121;
Matches 341; Conservative 64; Mismatches 149; Indels 179; Gaps 27;

13 GNNHNMWLGSLSPHKMEATSAATVPTT-----F 41
5 GGSNMWLGSLSPHKMEATSAATVPTT-----F 64
42 YMSPGSHL-----SNFGMCYGVEN--GNFHSPLTYMPLKSDGSLCTLALKRSQTV 92
65 LPSPPAAQVAVAPSGYYVYGAYGDTSTAGVYISHLPVMPKSDGSLCTING----- 117
93 VAVPTSSPYLEDFLGATWGTGHEYSGHERGLSDSIYNSQAEAPNNDLSQEPFROG 152
118 -NMPSSPKLEDPLFGNGSGH-----DATYVSGQEAED-----ASRAAYOH 160
153 HNSVQTHPYRSLACHGLYQALEBETTKETHTVSDSSLMPTWTEGLKRWVAPTRFESH 212
161 HOLV---PYN-----YQPLTEAMIQEAAAAPVEDAM-----AAKNF--- 195
213 QCVLEQQNMCGMNERNGVSLSGVGGEL-QSLISLMSPGSQ--SSCVT-APSGTDSVAVD 269
196 ---LVTSYACVGNQ-----EMPOSLISMSGSGSSCVSAAPQHQOQMAV 240
270 A-----KRGAAKLQKQPVHKSIDTRGQRTSQRGTVRRHW 307
241 AAAAAAGDQGSNSNDGEGORVGKRGKGTGQKQPVHKSIDTRGQRTSQRGTVRRHW 300
308 TGRYEALHMDNCKEKGOTRKGROYLYLGYDMEKKARAYDALAKYWGSPSTHINSIN 367
301 TGRYEALHMDNCKEKGOTRKGROYLYLGYDMEKKARAYDALAKYWGSPSTHINPLEN 360
368 YQVLEEMQNRQEVVAHLRRKSSGFSRGASLYRGVTRHGHGRQWQARIGRVAGKNDLY 427
361 YRDEIEEMERMTQEVVAHLRRSSGFSRGASLYRGVTRHGHGRQWQARIGRVAGKNDLY 420
428 LGFSTQEEAAAYDVAIKFRGANAVTNFDISRYVERIMASSNLAGELARRKX----- 483
421 LGFSTQEEAAAYDVAIKFRGANAVTNFDIRYDVERIMASSNLAGELARRKX----- 480
484 --DNDRPNKDIDYKSVTVSVNNEETVQVQAGNNNNNDSEMKVLFNHPGQQQQA----- 537
481 APDHVFLIGREL-----GATEEASAAATVTVG-----TWRMVV--HGSOQQQQAAC 523
538 NGNGSQOKIMNCNGYRNSAFSMALODLIGIDSVGQNMDESSKI-THRSNTSGLVT 596
524 EATADLQK-----GEFGDAHS-ALHGTIVGDVESAADETIDVPGGKISGINSNSSLV 577
597 SLSSSEASPEKRGPSLLFPMPMETKINPLIGTSVTS---MLPSPVQO---MRDSPA 649
578 SLNSSESGSERGLAMLVAKI-----HPTAVSLAAMPVMPAPAAAHVBPSPAI 630
QY 650 SLSHLPVYFASMTD 662
DB 631 --AHLVVFAMMTD 641

```

RESULT 4  
 ABB79639  
 ID ABB79639 standard; protein; 669 AA.  
 AC ABB79639;  
 DT 21-OCT-2002 (first entry)  
 DE Rice AINTEGUMENTA-like polypeptide OsANT2.  
 KW AINTEGUMENTA; ANT-like polypeptide; OsANT2; rice; transgenic plant;  
 plant.  
 OS Oryza sativa.  
 PN WO200259332-A2.  
 PD 01-AUG-2002.  
 PF 19-DEC-2001; 2001WO-US049294.  
 PR 21-DEC-2000; 2000US-0257896P.  
 PA (MONS) MONSANTO TECHNOLOGY LLC.  
 PI He SS, Dotson SB;  
 DR WPI; 2002-599798/64.  
 DR N-PSDB; ABBN84483, ABBN84484.  
 XX  
 PT New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in  
 improving agronomic, horticultural, and quality traits of plants, such as  
 increased size of plant organs.  
 PS Claim 3; Page 154-157; 169pp; English.  
 XX  
 CC The present sequence is the protein sequence of OsANT2, a newly  
 identified AINTEGUMENTA-like (ANT-like) polypeptide of rice. The sequence  
 was predicted from a genomic DNA sequence (see ABBN84483) obtained from a  
 rice database screening. An additional rice ANT-like polypeptide, OsANT1  
 (see ABB79638), was also identified. The invention provides nucleic acids  
 encoding ANT-like polypeptides comprising, in the N-terminal to C-  
 terminal direction, 2 AP2 DNA binding domains followed by an amino acid  
 subsequence selected from those given in ABB79629-35. ANT-like  
 polypeptides were identified in soybean, rice, cotton and corn (see  
 ABB79636-41). Nucleic acids (see ABBN84480-86) encoding the ANT-like  
 polypeptides can be used in the construction of transgenic plants,  
 especially corn, soybean, canola, wheat, cotton, tomato or potato (all  
 claimed) having improved agronomic, horticultural or other quality  
 traits, such as increased size of plant organs. These plants are  
 especially useful for production of ethanol or animal feedstuff

XX  
 SQ Sequence 669 AA;  
 XX  
 Query Match 38.5%; Score 1349.5; DB 5; Length 669;  
 Best Local Similarity 45.3%; Pred. No. 8e-116;  
 Matches 341; Conservative 66; Mismatches 142; Indels 203; Gaps 34;

QY 13 GNNHNLGFSLSPPMKNEAT-----SAATVPTTFYNSP 45  
 DB 18 GGVGGMWIGFSLSPHMAITYCAGVDVGHHHHHHHVHQOQHGGGLFYNPAAVAASFFYGG 77  
 QY 46 SQ-----SHLSNFGACVGVGENGNFHSPLTYPLKSDGSLCILEALK--RSOTQVMVPTSS 99  
 DB 78 GHDAVVTSAAGGSYTAG-----FSMPLKSGSLICIMALGSGQEGQGVVNS 129  
 QY 100 PKLEDFLG-GATNGTHEYSGHERGLSD--STYYNQNAEAQPNRLLSQPRQGHM-- 154  
 DB 130 PKLEDFLGAGPAM-----ALSLDNSAFYGGHC-----HHQGHAD 165  
 QY 155 --SVQTHPYYSG--LACHGLYQAPLEERTTKETHVSDGSLMPQMTGKQWVAPTRRF 209  
 DB 166 GGAVGSDPHHGGGSLQD-----AVIPGAGG----- 192

QY 210 STHQVLEQDNMC-----GMGNERNVSLGSVGC-----GETQSLSLSM- 248  
 DB 193 --HDAALVHDQSAALVAAGMAAHGGGYDINAAADVD--CAAGPIIPFGHHPITLSMS 249  
 QY 249 SPQSQSCVY--APSGTDSVAVD--KRGHA-KLGQKQVYHRKSIITFGQRTSQYGV 302  
 DB 250 SAGSQSCVYQAAAAAGPEYMAVDVSKRGADACQKQPVHRKSIDTFGQRTSQYGV 309  
 QY 303 TTRRTGRYEAHLNNSCKEGQTKRGQVYIGVDMEEKARAYDLALKYWGSPSTHN 362  
 DB 310 TTRRTGRYEAHLNNSCKEGQTKRGQVYIGVDMEEKARAYDLALKYWGSPSTHN 369  
 QY 363 FSIENYQVLEEMKNSRQEVYVAHLRRKSGFSRGNSTYRGVTRHGHGRWQARIQYAG 422  
 DB 370 FLEDYQVELEEMKNSRQEVYVAHLRRKSGFSRGNSTYRGVTRHGHGRWQARIQYAG 429  
 QY 423 NKDLVIGTF-----STOEBAEAYDVAATKRGANAVTNPDISRYVERIM 468  
 DB 430 NKDLVIGTFIAGAFAPARRARHAGTQEEAEAYDVAALFRGLAVTNFDITRYDVDKIL 489  
 QY 469 AASNLAGEELARRK-KDNDPRNKDIDYKSVTSYNNERTVOYQGNNNENDSEMKOYL 527  
 DB 490 ESTTLPGSLARRKKGVDGGG-----AAAVADAALVQAG-----NVAEMKAT 535  
 QY 528 ----FNHPSQQQANGNSDQKINCGNYRNSAFSMALQDLIG-IDVSGSGHMLDES 581  
 DB 536 AALPMAARTEDQOQHGHGHC--HHDLPLSDAFSV-LQDIVSTVDAAGAPPR----- 585  
 QY 582 SKIGTFHSNTSLVTSLSRSREASPEKRG-----SLFPMPMERTKIYNPIGTSV 632  
 DB 586 --APMMSMA--TSLGNSRREGSP-REVGGGGGGVLAATLPAKPAASXLYSPV--PL 636  
 QY 633 TSWL-PSPTVQMPSPA-ISLSHLPVFASWTD 662  
 DB 637 NTWASPSPAVSVPARAGVSIHLMPFAWTD 668

RESULT 5  
 AAB07724  
 ID AAB07724 standard; protein; 555 AA.  
 AC AAB07724;  
 DT 07-NOV-2000 (first entry)  
 DE An Arabidopsis aintegumenta (ANT) polypeptide.  
 KW Aintegumenta; ANT; cell proliferation; growth; organ mass; fertility;  
 asexual reproduction; plant; male sterile plant; female sterile plant;  
 early flowering.  
 OS Arabidopsis sp.  
 PN WO200040694-A2.  
 PD 13-JUL-2000.  
 PF 07-JAN-2000; 2000WO-US000465.  
 PR 08-JAN-1999; 99US-00227421.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Fischer RL, Mizukami Y;  
 DR WPI; 2000-465969/40.

Key Location/Qualifiers  
 Domain 281..357  
 FT /note="AP2 domain"  
 FT 383..451  
 FT /note="AP2 domain"

DR N-PSDB; AAA59220.  
 XX Modulating growth and cell proliferation in a plant used to alter organ  
 PT mass, control fertility and enhance asexual reproduction in plants  
 PT comprises modulating ANT activity and selecting plants with altered cell  
 PT number.  
 XX  
 XX  
 PS Claim 11, Page 41, 54pp; English.  
 XX  
 XX The present sequence represents an aintegumenta (ANT) polypeptide. The  
 CC ANT gene is expressed and functions not only in developing ovules but  
 CC also in various developing organs. Growth and cell proliferation in plant  
 CC can be modulated by modulating ANT activity. Modulation of ANT activity  
 CC is used to alter organ mass, control fertility and enhance asexual  
 CC reproduction in plants. Increased ANT activity can be used to produce  
 CC male or female sterile plants. Inhibition of ANT activity can be used to  
 CC truncate vegetative growth, resulting in early flowering  
 CC  
 XX  
 SQ Sequence 555 AA;

Query Match 37.9%; Score 1328; DB 3; Length 555;  
 Best Local Similarity 43.5%; Pred. No. 5.8e-114; Indels 226; Gaps 22;  
 Matches 314; Conservative 69; Mismatches 113;

QY 1 MRINESNTDDGNHNLGFSLSPH-MKM-----EATSAATVPTTFM 43  
 DB 1 MAFPCDNDNNHNTNLGLFSLSNMKGREGRAIYSSSTSAATSSSVPLQV 60  
 QY 44 SPSSGHLNFMGCGYGVGENGNHSPLTMPKSDSLCILALKRSQGVNVPSSPKLE 103  
 DB 61 G---DNTSNFGYCYGNPNPGIYSHMSVPLRSDSLCLMBALNRSSHNHSDSSPEVE 117  
 QY 104 DLGLGATGCTHYGSHERG--LSLSIYNSQNAEOPNDLSOPF-----ROGSHMSVQ 157  
 DB 118 DFFG---THNNTTHKEMDLSDLSLFYNTH---EENITTNQEFSPQTNNH---- 166  
 QY 158 THPYYSGLACHGLYQALEEFTTKETHVSDGSLMPQNTBGLKNWVAETREFTTHQVLE 217  
 DB 167 -----EEET--RNYGND-----PSLTHG----- 182  
 QY 218 QQMNCMGNERNGVSLGSGCEL--QSLSLMSPGSSSCVTAPs----- 261  
 DB 183 -----GSFNVG-----YEFQOSSLMSPGSSSCITGSHHQNQONHQSNNH 229  
 QY 262 -----GTDVAVDAKRGH---AKLGQKOPVHRKSIDTFGORTSOYRGVTR 304  
 DB 230 QQISEALVETSVGFETTWAAKKRGQEDVYVVOQKQIVHRKSIDTFGORTSOYRGVTR 289  
 QY 305 HMTGRTYEAHMDNCKECCGCTRKROYVLLGGYDMEKAAARAYDLAAKYGSPSTHNFs 364  
 DB 290 HMTGRTYEAHMDNCKECCGCTRKROYVLLGGYDMEKAAARAYDLAAKYGSPSTHNFs 349  
 QY 365 IENYQVLEEMKMKNSROEYVAHLRRKSSGFSRNASIYGVTRRHQHGWRQARIGRVAGNK 424  
 DB 350 AENYQVLEEMKMKNSROEYVAHLRRKSSGFSRNASIYGVTRRHQHGWRQARIGRVAGNK 409  
 QY 425 DLYLGFSTQGEAAEAYDVAALKFRGANAVTNFDISRDVVERIYASNLGELARRKCD 484  
 DB 410 DLYLGFSTQGEAAEAYDVAALKFRGANAVTNFDISRDVVERIYASNLGELARRKCD 467  
 QY 485 NDPRNKDIDYNSVVTSVNNEETVOYQAGNNNNENDSEMKVLFNHPSSQOQOANGNSDQ 544  
 DB 468 -----NNSIV--VANTE-----DQ 479  
 QY 545 KIMNCNRYNSAFSMAQLDLIGDIVSGCHMLDESSKIGTHFENTSSIVTSLSSREA 604  
 DB 480 TALNA-----VEGSSNREV 494  
 QY 605 SPEKRGPSL--LFPMPMETKIV-NEIGTSVTSWLPSPVQWRPSPALSLHLPVFASWT 661  
 DB 495 STPERLSTFPAIFALQOVNOKKFGSMNGSMSPMISNPAELK--IVALLTLPMQPFALMA 553  
 QY 662 DT 663

DB 554 DS 555  
 RESULT 6  
 AAG31380  
 ID AAG31380 standard; protein; 555 AA.  
 XX  
 AC AAG31380;  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 37675.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139751P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 28-JUN-1999; 99US-0140695P.  
PR 29-JUN-1999; 99US-0140823P.  
PR 30-JUN-1999; 99US-0140991P.  
PR 01-JUL-1999; 99US-0141287P.  
PR 02-JUL-1999; 99US-0141544P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142920P.  
PR 13-JUL-1999; 99US-0142977P.  
PR 14-JUL-1999; 99US-0143542P.  
PR 15-JUL-1999; 99US-0143624P.  
PR 16-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144633P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145090P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 25-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148644P.  
PR 16-AUG-1999; 99US-0149368P.

PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151308P.  
PR 01-SEP-1999; 99US-0151309P.  
PR 07-SEP-1999; 99US-0152633P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 14-OCT-1999; 99US-0159684P.  
PR 18-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 25-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.9%; Score 1328; DB 3; Length 555;  
Best Local Similarity 43.5%; Pred. No. 5.8e-114;  
Matches 314; Conservative 6; Mismatches 113; Indels 226; Gaps 22;

QY 1 MKRINSENNITDGGNNHNLGFSLSPH-MKN-----EATSAATVPTTFYM 43  
DB 1 MKSFCDNDNNHNSVTTLLGFSLSNNMKKGGKGRALVSSSTSSAATSSSVPPOLVV 60  
QY 44 SPQSGHLSNFGMKCYGVENENFHSPLTWPLKSGSGSLICLALKRSGQVWVPTSSPKLE 103  
DB 61 G--DNTSNFGVCYGSNPNCGTISHMSVPLRSDGSLCLMEALNRSSHSYHHODSSPKVE 117

QY 104 DFLGATMGTHEYSGHERG--LSLDSIYNSQNAEAPNDLSQPF-----ROQGHMSVQ 157  
 DB 118 DFFG-----THHNTSHKEMDLSDLSLFYNTTH---EPNTTNQGFESFPQTRNH----- 166  
 QY 158 THPYSGLACHGLYQAPLEETTKETHVSDCCSLMPQMTBGLKNWVAPRFEFSTHQVLE 217  
 DB 167 -----EERT--RNYGND-----PSLTHG----- 182  
 QY 218 QQNMCGMNERNGVSLGSGGCEL-QSLSLSMSPGSGSSCVTAPS----- 261  
 DB 183 -----GSFNVGV-----YGEFQOSLSLSMSPGSGSSCVTSHHHQONQNHQSONH 229  
 QY 262 -----GTDVAVDAKKRGH---AKLGKQPYHRKSIDTFGQRTSQYRGVTR 304  
 DB 220 QQISEALVETSVGFETTTMAAAKKRGQEDVWVVGQKQIVHRKSIDTFGQRTSQYRGVTR 289  
 QY 305 HRMTGRYEAHLMDNSCKKEGCTRKGRQVYLGVDMEKARAAYDLAAKTMGSPSTHNF 364  
 DB 230 HRMTGRYEAHLMDNSFKKEGHSRKGQVYLGVDMEKARAAYDLAAKTMGSPSTHNF 349  
 QY 365 IENYQVLEEMKMSRQEVYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQARIGRVAGNK 424  
 DB 350 AENYQKEIEDMKMTRQEVYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQARIGRVAGNK 409  
 QY 425 DLYGTSTGTEBAAEAAYVAALKFRGANAAVTNPDISYDVERIMASSNLLAGELARRKD 484  
 DB 410 DLYGTSTGTEBAAEAAYVAALKFRGINAVTNPDITRIVDRIMSSNTLLSGELARRN-- 467  
 QY 485 NDPNRKQIDYKSVVTSVNNSEETVQVQAGNNNNENDSEMKVLFNHP90QOQANGNSDQ 544  
 DB 468 -----NNSIV--VRNTE-----DQ 479  
 QY 545 KIMNGCNRNSAFSMALODLIGDSVSGQHMLDESKIGTHSNTSLTSLSSSREA 604  
 DB 480 TALNA-----VEGSKKEV 494  
 QY 605 SPEKQSPSL--LEPMPMETKIV-NPIGSVTSMSPPTVQMRSPATISLHLPVFASWT 661  
 DB 495 STPEELSLFPALFALPQVQKMFSGNSMGNSPMTSNPAELK-TVALLTLQMTVFAPAMA 553  
 QY 662 DT 663  
 DB 554 DS 555  
 RESULT 7  
 ABR40832  
 ID ABR40832 standard; protein; 555 AA.  
 XX ABR40832;  
 AC  
 XX  
 DT 16-MAY-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana oil trait related protein sequence SEQ ID NO:424.  
 XX  
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 KM receptor-like protein kinase; mltogen activated protein kinase; oil;  
 KM lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 KM CMC-like transcription factor; antisense inhibition; co-suppression;  
 KM transgenic plant.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO2003002751-A2.  
 XX  
 PD 09-JAN-2003.  
 XX  
 XX 27-JUN-2002; 2002WO-US020152.  
 PF  
 XX 29-JUN-2001; 2001US-0301913P.  
 PR  
 XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
 PA (PION-) PIONEER HI-BRED INT INC.

XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harwell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczy-Haski MC,  
 DR WPI, 2003-201509/19.  
 XX  
 PT Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity, caleosin-like activity, useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 PS Claim 12; Page 463-465; 542pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (pe) having receptor-like protein kinase activity, mltogen  
 CC activated protein (MAP)-kinase activity, lipid-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity, and CMC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV) or  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00686 and ABR40591 to  
 CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 555 AA;  
 Query Match 37.9%; Score 1328; DB 6; Length 555;  
 Best Local Similarity 43.5%; Pred. No. 5,8e-114;  
 Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;  
 QY 1 MKRINSNTTDDGNNTNMLGFSLSPE-MKM-----EATSAATVPTTFYM 43  
 DB 1 MKSFCDNDNNHNTNTNLLGFLSSMMMGAGRGREAIYSSSTCSAATSSSVPOLVV 60  
 QY 44 SPQSHLSNFGMCYGVGNGNPHSPITVPLPKSDGSLCILEAKSQOTVMPPTSPKE 103  
 DB 61 G---DNTSNFGVGYGNGNPGIGTSHMSVPLSDGSLCMEALNRSNHHQDSSPKYE 117  
 QY 104 DFLGATMGTHEYSGHERG--LSLDSIYNSQNAEAPNDLSQPF-----ROQGHMSVQ 157  
 DB 118 DFFG-----THHNTSHKEMDLSDLSLFYNTTH---EPNTTNQGFESFPQTRNH----- 166  
 QY 158 THPYSGLACHGLYQAPLEETTKETHVSDCCSLMPQMTBGLKNWVAPRFEFSTHQVLE 217  
 DB 167 -----EERT--RNYGND-----PSLTHG----- 182  
 QY 218 QQNMCGMNERNGVSLGSGGCEL-QSLSLSMSPGSGSSCVTAPS----- 261  
 DB 183 -----GSFNVGV-----YGEFQOSLSLSMSPGSGSSCVTSHHHQONQNHQSONH 229  
 QY 262 -----GTDVAVDAKKRGH---AKLGKQPYHRKSIDTFGQRTSQYRGVTR 304  
 DB 220 QQISEALVETSVGFETTTMAAAKKRGQEDVWVVGQKQIVHRKSIDTFGQRTSQYRGVTR 289  
 QY 305 HRMTGRYEAHLMDNSCKKEGCTRKGRQVYLGVDMEKARAAYDLAAKTMGSPSTHNF 364  
 DB 230 HRMTGRYEAHLMDNSFKKEGHSRKGQVYLGVDMEKARAAYDLAAKTMGSPSTHNF 349  
 QY 365 IENYQVLEEMKMSRQEVYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQARIGRVAGNK 424  
 DB 350 AENYQKEIEDMKMTRQEVYVAHLRRKSSGFSRGASIRYGVTRHHQGRMQARIGRVAGNK 409  
 QY 425 DLYGTSTGTEBAAEAAYVAALKFRGANAAVTNPDISYDVERIMASSNLLAGELARRKD 484  
 DB 410 DLYGTSTGTEBAAEAAYVAALKFRGINAVTNPDITRIVDRIMSSNTLLSGELARRN-- 467



KW asexual reproduction; plant; male sterile plant; female sterile plant;  
 early flowering.  
 XX Brassica sp.  
 XX WC2000040694-A2.  
 XX 13-JUL-2000.  
 XX 07-JAN-2000; 2000MO-US0000465.  
 XX 08-JAN-1999; 99US-00227421.  
 XX (REGC ) UNIV CALIFORNIA.  
 XX PA  
 XX Fischer RL, Mizukami Y;  
 XX WPI; 2000-465969/40.  
 XX N-PSDB; AAA59222.  
 XX  
 XX Modulating growth and cell proliferation in a plant used to alter organ  
 mass, control fertility and enhance asexual reproduction in plants  
 PT comprises modulating ANT activity and selecting plants with altered cell  
 number.  
 XX  
 XX Claim 15; Page 44; 54pp; English.  
 XX  
 XX The present sequence represents an aintegumenta (ANT) polypeptide. The  
 CC ANT gene is expressed and functions not only in developing ovules but  
 CC also in various developing organs. Growth and cell proliferation in plant  
 CC can be modulated by modulating ANT activity. Modulation of ANT activity  
 CC is used to alter organ mass, control fertility and enhance asexual  
 CC reproduction in plants. Increased ANT activity can be used to produce  
 CC male or female sterile plants. Inhibition of ANT activity can be used to  
 CC truncate vegetative growth, resulting in early flowering  
 CC  
 XX Sequence 548 AA;  
 XX  
 XX Query Match 37.3%; Score 1307; DB 3; Length 548;  
 XX Best Local Similarity 43.6%; Pred. No. 5.1e-112;  
 XX Matches 309; Conservative 77; Mismatches 103; Indels 220; Gaps 25;  
 QY 9 NTDDGNNHNVWLGFSLSPEH-----KQKATSAATVPTFYMSPSQ-----SHLSNFGWCY 57  
 DB 6 DDDDDNTTLDFSSLSNMLKMGSGEALYSSSSSVATS--SVPRQLVYGDNSNXYGCY 63  
 QY 58 GVG-ENGFFHSPPLTWMPKSDGSLCTLEAKRS-----QTVWVPTSPFKLEDFLGA 109  
 DB 64 GSNLAAREMYSQMSVPLRSDSLCIMEALNRSNHHNSOV---SSPKMEDF----- 115  
 QY 110 TNGTEHYG-SHERG--LSLDSLYNSONAEOQPNDDLQPPROGCHMSVOTHYYSGLA 166  
 DB 116 -PGTHHNTSHKANDLSLDFYNTTHA--PNNNTVQEF---EFPTQRNHH----- 163  
 QY 167 CHGLYQAPLEERTYKTHVSDCSSLMPQTEGLKXWVAPTREFSTHQVLEQMNCGMN 226  
 DB 164 -----EERT--RYNEND-----PGLTHG----- 179  
 QY 227 ERNGVSLGSGGEL-QSLSLSMSPGSSSCVTA-----P 260  
 DB 180 ---GSGFNVYGVGEFQOSLSLSMSPGSSSCITASHHONOTONHOQISEALVETSAGFE 236  
 QY 261 SGTDSVAVDAKRRG-HAKLGQKQPYHRKSIDTFGQRTSQYRGVTRHRTGRYEALWMS 319  
 DB 237 TTMTMAAAKKRRGQGVVVGQKQIVHRKSIDTFGQRTSQYRGVTRHRTGRYEALWMS 296  
 QY 320 CKKEGQTRKQRYLYGTYMEKARAYDLALKTKWGSTHNFSEIENYVQVLEMKMS 379  
 DB 297 FKKEGSRKGRQYLYGTYMEKAAAYDLALKTKWGSTHNFSEIENYVQVLEMKMS 356  
 QY 380 ROEYVHLRRKSGSGFRGASIRYGVTRHGHGRWQARIGRVAGNDLVLTGFTQEEAAE 439  
 DB 357 ROEYVHLRRKSGSGFRGASIRYGVTRHGHGRWQARIGRVAGNDLVLTGFTQEEAAE 416

QY 440 AYDVAIKFRGANAVTNFDISRYDERIMASSNLLAGELARKKKNDPRNKDIDYKSV 499  
 DB 417 AYDVAIKFRGNATNTFDITRYDVRIMASNTLLSGEMARRNS-----NSIVV 465  
 QY 500 TSVNNEE---TVQVQAQNNNNENSEMKVLPFNHESQQQQAANGNSDDKIMKCGYRNSA 556  
 DB 466 RNISDEEALTVAVNGSKREV-----GSPERVLSF----- 496  
 QY 557 FPMALQDLDIGDSVSGGHNMLDESSKIGTHESNTSLVTSLSRSSREASPEKGPILLFP 616  
 DB 497 -----PTIFA 501  
 QY 617 MPPMETKI--VNPIGTSVTSWLPSPVQMRSPPAISLSHLVYFASWTD 663  
 DB 502 LPQVGPKNFGANVVG-NMSSWTNNADLK-TVSLTLQMPVFAAMADS 548  
 RESULT 10  
 AAG31382  
 ID AAG31382 standard; protein; 528 AA.  
 XX  
 XX AAG31382;  
 AC  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 37677.  
 XX  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 XX hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX EPI033405-A2.  
 XX  
 XX PD  
 XX 06-SEP-2000.  
 XX  
 XX PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 XX 25-FEB-1999; 99US-0128125P.  
 XX 05-MAR-1999; 99US-0123180P.  
 XX 09-MAR-1999; 99US-0123548P.  
 XX 23-MAR-1999; 99US-0125788P.  
 XX 25-MAR-1999; 99US-0126264P.  
 XX 29-MAR-1999; 99US-0126785P.  
 XX 01-APR-1999; 99US-0127462P.  
 XX 06-APR-1999; 99US-0128234P.  
 XX 08-APR-1999; 99US-0128714P.  
 XX 16-APR-1999; 99US-0129845P.  
 XX 19-APR-1999; 99US-0130077P.  
 XX 21-APR-1999; 99US-0130449P.  
 XX 23-APR-1999; 99US-0130510P.  
 XX 23-APR-1999; 99US-0130891P.  
 XX 28-APR-1999; 99US-0131449P.  
 XX 30-APR-1999; 99US-0132048P.  
 XX 30-APR-1999; 99US-0132407P.  
 XX 04-MAY-1999; 99US-0132484P.  
 XX 05-MAY-1999; 99US-0132485P.  
 XX 06-MAY-1999; 99US-0132486P.  
 XX 06-MAY-1999; 99US-0132487P.  
 XX 07-MAY-1999; 99US-0132863P.  
 XX 11-MAY-1999; 99US-0134256P.  
 XX 14-MAY-1999; 99US-0134218P.  
 XX 14-MAY-1999; 99US-0134219P.  
 XX 14-MAY-1999; 99US-0134221P.  
 XX 14-MAY-1999; 99US-0134370P.  
 XX 16-MAY-1999; 99US-0134768P.  
 XX 15-MAY-1999; 99US-0134841P.  
 XX 20-MAY-1999; 99US-0135124P.  
 XX 21-MAY-1999; 99US-0135353P.  
 XX 24-MAY-1999; 99US-0135629P.  
 XX 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140355P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142300P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147303P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 18-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149266P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151666P.  
PR 27-AUG-1999; 99US-0151800P.  
PR 30-AUG-1999; 99US-0151903P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 11-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155559P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 04-OCT-1999; 99US-0157153P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158322P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159243P.  
PR 13-OCT-1999; 99US-0159244P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 13-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160800P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.





```
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144312P.
PR 19-JUL-1999; 99US-0144312P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 26-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149425P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149922P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.0%; Score 1295.5; DB 3; Length 529;
Best Local Similarity 44.6%; Pred. No. 5,7e-111;
Matches 301; Conservative 63; Mismatches 102; Indels 205; Gaps 20;

QY 31 ATSAATVPTFTTNSPQSHLSNFGMCYGVGENGNFSP:LTWPLKSDGSLCTLEAIKRSQ 90
Db 22 ATSSSVPPQLVYG--DNTSNFGVCYGSNPGIYSHSNVPLRSDGSLCTLEAIKRRSS 78
QY 91 TCVMVPTSPKLEDFLGCATMGTHGESHGRC--LSLDSIYNSQNAEAPRDLISQPF 148
Db 79 HSNHODSSPKVEDDFG---THNNTSHKEMDLSLDSLFYNTH--EPVTTNFDQF 131
QY 149 ---RQGHMSVQTHRYSGLAGHLYOAPLEETFKETHVSDGSLMPQMTGLKNWYA 204
Db 132 FSPQTRNH-----FEET--RNYGND-----PSLTHG----- 156
QY 205 PTFREFTHQOVLEQONCGMGNERNKVSLSGVCCEL--OSLSLSMSPGSSCVYAP.S-- 261
Db 157 -----GSFNVGV-----YGFQOSSLNSPQSSCLITGSHHH 190
QY 262 -----GTDVAVDAAKRGH--AKGQKQPVHRKSIDT 291
Db 191 QONONHOSONHQISEALVETSVGFETTMAAKKRGQDQVYVVGQKQVHRKSIDT 250
QY 292 FQGRTSQVGRVRRHRTGTGYEHLMDNSCKEGQVRKGOVYLGVDMEKRAAYDLAA 351
Db 251 FQGRTSQVGRVRRHRTGTGYEHLMDNSFKEGHRSKGRQVYLGVDMEKRAAYDLAA 310
QY 352 LKTMGPSTHINFSIENYQVLEEMKMSROEYVAHLRRKSGFSRQASIRGVTHRHQHG 411
Db 311 LKTMGPSTHINFSIENYQVLEEMKMSROEYVAHLRRKSGFSRQASIRGVTHRHQHG 370
QY 412 RMOARIGRYAANKDYLGTFTQEBAAEAYDVAAIKFGCAAVTTFDISRIVERIMASS 471
```

Db 371 RMOARIGRVAGNKDLYGTGTGTOEAAEAVDAAIKRGNTAVNTNPDITRYDVDRIMSN 430  
 QY 472 NLAGEIARRKKNDPRNKIDYKNSVYTSVNEETVQVQAGNNNNENDEMKVLFNHP 531  
 Db 431 TLIGELARRR-----NNSIV--VRNTE-----451  
 QY 532 SQQQOANGNSDQKIMNCNTRNSAFSMALODLIGDVSQGGHNMIDESSKIGTHRSNT 591  
 Db 452 -----DQALNA-----458  
 QY 592 SSVYTSLSRSEAPSEKRGPSL--LFPMPMETKIV-NPIGTSVTSWLPSPVQMRPSPA 648  
 Db 459 ---VVEGSSKKEVSTPEHLSPFPAIFALPQVNOGMFGSNMGNNSPMTSPNAELK-TVA 514  
 QY 649 ISLSHLPVFASWTD 663  
 Db 515 LTLPMQPVFAAMADS 529

RESULT 12  
 ID ABB79640 standard; protein; 585 AA.  
 AC ABB79640;  
 DT 21-OCT-2002 (first entry)  
 DE Cotton AINTEGUMENTA-like polypeptide GhANT1.  
 KW AINTEGUMENTA; ANT-like polypeptide; GhANT1; transgenic plant; cotton;  
 KM plant.  
 OS Gossypium hirsutum.  
 PN WO200259332-A2.  
 PD 01-AUG-2002.  
 PF 19-DEC-2001; 2001WO-US049294.  
 PR 21-DEC-2000; 2000US-0257896P.  
 PA (MONS ) MONSANTO TECHNOLOGY LLC.  
 PI He SS, Dotson SB;  
 DR WPI, 2002-599798/64.  
 XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in  
 PT improving agronomic, horticultural, and quality traits of plants, such as  
 PT increased size of plant organs.  
 PS Claim 3: Page 158-161; 169pp; English.

The present sequence is the protein sequence of GhANT1, a newly  
 identified AINTEGUMENTA-like (ANT-like) polypeptide of cotton. The  
 sequence was predicted from a cDNA clone (see ABB84485) isolated in a  
 cotton DNA database screening. The invention provides nucleic acids  
 encoding ANT-like polypeptides comprising, in N-terminal to C-terminal  
 direction, 2 AP2 DNA binding domains followed by an amino acid  
 subsequence selected from those given in ABB79629-35. ANT-like  
 polypeptides were identified in soybean, rice, cotton and corn (see  
 ABB79636-41). Nucleic acids (see ABB84480-86) encoding the ANT-like  
 polypeptides can be used in the construction of transgenic plants,  
 especially corn, soybean, canola, wheat, cotton, tomato or potato (all  
 claimed) having improved agronomic, horticultural or other quality  
 traits, such as increased size of plant organs. These plants are  
 especially useful for production of ethanol or animal feedstuff

Sequence 585 AA;  
 Query Match 33.7%; Score 1180.5; DB 5; Length 585;  
 Best Local Similarity 41.7%; Pred. No. 3.4e-100;

Matches 290; Conservative 79; Mismatches 164; Indels 163; Gaps 25;  
 QY 17 NWLGSLSPHMKMEATSAATVPTTFYMSPGSHLSFMCMCVGNGNFGHSPITMPLKS 76  
 Db 3 NWLGSLTFPDLRIDSS-----FREDHGFPSS--VMPURS 35  
 QY 77 DGLSLIEALKRSQR-----QVWPTSSPKLEDEPLG----GATGTHEY-GS 118  
 Db 36 DGLSLCVDEPFRRSSTIADEDMRYENGIGSATINEGPKLEDEPLGCGYSNPSQETAYCGT 95  
 QY 119 HER-----GLSDSIYNSQNAEAQPN-----RDLSQP--FROQGNMS 155  
 Db 96 HENQNTVPSPTRINNVAVAPNYSSGDAEAEENFTNPSSFIQTRMYNENPOTLMAGHSL 155  
 QY 156 VQTHPYVSLACHGLYQAPLEETTKETHTVSDSSSLMPQMTGLKNWVAFTPEFTHQCV 215  
 Db 156 QQCDPVPNNQSGVHHVDPESATS-----VSGFSWLRLQT-PFPGK-- 197  
 QY 216 LEQWNCGMENRNGVSLSGVCGELQSLSLMSPSQSSCVTAPSGTDSVA---VDK 271  
 Db 198 -----ASGNETNNNF-----NFOALSLTMSF-----TSRNGFPALAPLEVVDNR 236  
 QY 272 KKGHALGQKQRYHRKSIDTFCQRTSOYRGVTRHRMTGTYEHLMDNSCKEGQTRKQ 331  
 Db 237 KRPVGNLITRESVPRKSIDTFCQRTSOYRGVTRHRMTGTYEHLMDNSCKEGQTRKQ 296  
 QY 332 VYLGYYDEEKARAYDLAALKYMGSPTHINFSIENYCYOLEEMKMSROEYVAHLRRKS 391  
 Db 297 VYLGYYDEEKAKAKVDLAALKYMGTHINPLSTYEKELEEMKMTQOEVAHLRRKS 356  
 QY 392 SGFSRGASLYRGVTRHHQGRWOARIGRYAGNKDLYLGFSTQEEAAEYDVAALKFRRA 451  
 Db 357 SGFSRGASVYRGVTRHHQGRWOARIGRYAGNKDLYLGFSTQEEAAEYDVAALKFRRT 416  
 QY 452 NATNFDIRYDVERIMASSNLLAGELARRKNDPRANDIDYKNSVTSVNEETVQVQ 511  
 Db 417 SATNFDISRYDYKRCSSSTLIGELARRSKPDASIAPEBYN-SCASSAPQPLAIP 475  
 QY 512 AGNNNNE-NDSEWKVYLFNHPSQQQ-QANGSGDQKIMNCNTRNSAFSMALODLIGT-- 567  
 Db 476 SGASDELADMTWTA---NSDEQQCHOSTNTNNDASLANSS-RUSSNPSQSRGSLGAS 531  
 QY 568 DSVSGQHNMIDESSKIGTHFSTSLVTSLSRSEAPSEKRGPSLFPMPMETKIVNP 627  
 Db 532 DKFGIG---GDYSHHG-YFSLKSGKYEDGNSSTINSNENR-----567  
 QY 628 IGTSVTSWLPSPVQMRPSPAISLSH-LPFAWTD 662  
 Db 568 LGN-----LGLVHKIMPFALMNE 585

RESULT 13  
 ID ABR40856 standard; protein; 557 AA.  
 AC ABR40856;  
 DT 16-MAY-2003 (first entry)  
 DE Glycine max oil trait related protein sequence SEQ ID NO:491.  
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 KM receptor-like protein kinase; mitogen activated protein kinase; oil;  
 KM lipid-like transcription factor calciosin; ATP citrate lyase; SNF1;  
 KM CLK-like transcription factor; antisense inhibition; co-suppression;  
 transgenic plant.  
 OS Glycine max.  
 PN WO2003002751-A2.  
 PD 09-JAN-2003.

PF 27-JUN-2002; 2002WO-US020152.  
 XX  
 PR 29-JUN-2001; 2001US-0301913P.  
 XX  
 XX (DUPO.) DU PONT DE NEMOURS & CO E. I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PA  
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harwell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczyński MC;  
 XX  
 DR WPI: 2003-201509/19.  
 DR N-PSDB; ACC00850.  
 XX  
 PT Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity, caleosin-like activity, useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 PS Claim 12; Page 509-511; 542pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
 CC activated protein (MAP)-kinase activity, lipid-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity and CRC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV); and (5)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to  
 CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention  
 CC  
 SQ Sequence 557 AA;

Query Match 30.3%; Score 1061.5; DB 6; Length 557;  
 Best Local Similarity 43.1%; Pred No. 3.8e-83;  
 Matches 260; Conservative 67; Mismatches 125; Indels 151; Gaps 23;

QY 100 PLELEPLGATGTHGTHGSHERGLSLDIYNSQNA-----BAQPNRDLLOQPFROQGM 154  
 |||||  
 DB 66 PKLEDFLG-----CYSDSPAKYFCQDSQPDQ---NQSQNNWSKI 101  
 QY 155 SVQTHYYVYSGLA CHGLYQAPLEETTKETHVSDCSLMPQW-----TEGL-KN 201  
 |||||  
 DB 102 NNVAVNSF-----CTN-----DSELETGNTLTPSSLSLSPHAYNDSHALPTNGMKXS 151  
 QY 202 WVAPEFSTHQVLEQWNCQGNERNNGVSLGVSQCGELQSLSLSPSGQSSCVTAP 261  
 |||||  
 DB 152 WLAQT-QFSQDGR-----PSNEANQCN-----FQSLSLTMSPSVQNG-VGALIS 192  
 QY 262 GTDGSVAVDAKKGHAALGOKOPVPHKRSITTEGQRTSOVRGVTGRHWTGRYEAHMDNCK 321  
 |||||  
 DB 193 SV-QVNEDSRKRYMAASHAREPVRKSIDTFQGRFSQVRGVRHAWTGTGYEHLMDNCKR 251  
 QY 322 KEGQTRKGRQVYLGADYMEKAKARAYDALALKYKWPSTHINSINYQVLEEMQMQRQ 381  
 |||||  
 DB 252 KEGQTRKGRQVYLGADYMEKAKAKAYDALALKYKWPSTHINSINYQVLEEMQMQRQ 311  
 QY 382 EYVAHNRKSSGSRGASIRYGVTRHNGRQWQARIGRVANKDYLIGTFSQGEAARAY 441  
 |||||  
 DB 312 EYVAHNRKSSGSRGASIRYGVTRHNGRQWQARIGRVANKDYLIGTFSQGEAARAY 371  
 QY 442 DVAALIFRGANAVTNDISRYVERIMASSNLAGEIARRKQNDPKNKI-DYN---KS 497  
 |||||  
 DB 372 DVAALIFRGANAVTNDISRYVERIMASSNLAGEIARRKQNDPKNKI-DYN---KS 431  
 QY 498 VVTSVNNBETVQVQAGNNNNE-NDSEWKNVLFNHSQCGQAGANGSDQKTNACNGYRNSA 556

DB 432 PMPVLSQPPPLAITDGHSHDELNNMM-----NANNSDQAPQ----- 469  
 QY 557 FSNMALODLIGIDVSGGQHNMLDESSKIGTGFHSNTSLVTSJSSSRERASPERKRGSLFP 616  
 DB 470 -----ESG--GAEFNNN---VTSSSSSQVSPSS-NKDALNP 500  
 QY 617 MPMETKIVNPIGTSTSW-----LPSPVQ-----MRSPAIS.SHLVPFAS 659  
 DB 501 QSP-----NEFGVSGADYGHGYFTLDGPKYDDGANNNDMSTRNLGNLGVNQVPMPL 554  
 QY 660 WTD 662  
 DB 555 WNE 557  
 RESULT 14  
 ID ABR40768  
 ID ABR40768 standard; protein; 710 AA.  
 XX  
 AC ABR40768;  
 XX  
 DT 16-MAY-2003 (first entry)  
 DE Glycine max oil trait related protein sequence SEQ ID NO:356.  
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 XX receptor-like protein kinase; mitogen activated protein kinase; oil;  
 XX lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 XX CRC-like transcription factor; antisense inhibition; co-suppression;  
 XX transgenic plant.  
 XX  
 OS Glycine max.  
 XX  
 XX W02003002751-A2.  
 XX  
 PD 09-JAN-2003.  
 XX  
 XX 27-JUN-2002; 2002WO-US020152.  
 XX  
 PF 29-JUN-2001; 2001US-0301913P.  
 XX  
 PR (DUPO.) DU PONT DE NEMOURS & CO E. I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PA  
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harwell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczyński MC;  
 XX  
 DR WPI: 2003-201509/19.  
 DR N-PSDB; ACC00803.  
 XX  
 PT Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity, caleosin-like activity, useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 PS Claim 12; Page 374-376; 542pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
 CC activated protein (MAP)-kinase activity, lipid-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity and CRC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to

CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 XX Sequence 710 AA;

Query Match 30.3%; Score 1059.5; DB 6; Length 710;  
 Best Local Similarity 36.8%; Pred. No. 8,8e-89;  
 Matches 284; Conservative 74; Mismatches 203; Indels 211; Gaps 24;

QY 17 NMLGSLSPKMKKATLS-AATVPTTFMSP-----SQSHLSNFGMXY 57  
 DB 5 NLLGSLSPKMKKATLS-AATVPTTFMSP-----SQSHLSNFGMXY 57  
 QY 58 GVGE-----NGNFSPLTVMPLKSD-----GSLCILEALKRSQTVWVPTSSPKLEDF 105  
 DB 64 GIVAFHRSNNIH---TTQDWKENYNSQNLGTSQNNQNNHQQQQ---QQQPLLENF 118  
 QY 106 LGGATMGTHGYGSHRGSLSDSIYNSQNAEQPNRDLSPFRQGGHMSVQTHPYYSGL 165  
 DB 119 LGG-----HSFGEHE-----QF-----YGGN 134  
 QY 166 ACHGLYQAPLEBEET-----TKETHVDCSSIMPQTEGLKXWVAPTRFSTHQVLEQ 218  
 DB 135 SASSTYMPAPQVLAGGGGGGSSNSNTSSISGLSM---IKTWL---RNPZPHSENNN 188  
 QY 219 QMNCGMGNERNGVSLGSGCGELQSLSLMSFGSSQSCCTAPSGTDSVAVDAKKRGHAKL 278  
 DB 189 NNNSSGGNSRESVO-----QTLISLMSGTGQS--TSLPLLTASV---DNSESS 233  
 QY 279 GQKQP-----VHKSIDTFGQRTSOVGVTRHRTYFAHLMDSCK 322  
 DB 234 DNKQPHHTALDTQTQGAIEAPKSIDTFGQRTSOVGVTRHRTYFAHLMDSCK 323  
 QY 323 EGQTRKGRQVYLGQYDMEKKAARAADALAKYWPSTHINFSLENYVQLENNKNSRQE 382  
 DB 294 EGQTRKGRQVYLGQYDMEKKAARAADALAKYWPSTHINFSLENYVQLENNKNSRQE 383  
 QY 383 YVAHLRRKSSGFSGASIRGVTHHGHQWQARIGVAGNKDLYLGTPTQGEAAEAYD 442  
 DB 354 YVASLRKSSGFSGASIRGVTHHGHQWQARIGVAGNKDLYLGTPTQGEAAEAYD 413  
 QY 443 VAAIKFGANAVTFDISRYDVERIMASSNLLAGELARRKKD-----NDPENKDIDY 494  
 DB 414 VAAIKFGANAVTFDISRYDVERIMASSNLLAGELARRKKD-----NDPENKDIDY 473  
 QY 495 NKSVVTS-----VNEETVOQAGNNNNENDSEWKKVLFNHPQOQOQANNGSDQKIM 548  
 DB 474 HSSIMNSHLLTQGIINN---YAAAGTTATHHHNNHNAALAFQOPCTTIHYPGQRINW 528  
 QY 549 CGNVR---NSAFPMALODLIGIDSVGSGQHMLDESSKIGTHFANTSSLYTSLSSREAS 605  
 DB 539 CKQEDNDSDASHSLYSIDHQLQGNNGTHNPFHTNSGLHMLMDASIDNSSSSSVV 588  
 QY 606 PEKGRPSLLFPMPMETKIV-----NP----- 627  
 DB 589 YDVGGGGGYVNIPIGTTTTVANDGQNPSSNHFQDNEIKALGYESVGGSTDPYHAH 648  
 QY 628 -----IGTSVTSMLPS--PTVQMPSPATSLSH 653  
 DB 649 ANLIYLLTQOQSSVDVAKAYDQSAQNTWVPAITTAHPRSTSMALCH 700

RESULT 15  
 ABR40766  
 ID ABR40766 standard; protein; 707 AA.

XX ABR40766;  
 XX  
 XX 16-MAY-2003 (first entry)  
 DT  
 XX  
 DE Glycine max oil trait related protein sequence SEQ ID NO:352.  
 XX  
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;

KW receptor-like protein kinase; mitogen activated protein kinase; oil;  
 KW LIPID-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 KW KCC-like transcription factor; antisense inhibition; co-suppression;  
 KW transgenic plant.

OS Glycine max.

XX WO2003002751-A2.

XX 09-JAN-2003.

XX 27-JUN-2002; 2002WO-US020152.

XX 29-JUN-2001; 2001US-0301913P.

XX (DUPO) DU PONT DE NEMOURS & CO B. I.

XX (PION-) PIONEER HI-BRED INT INC.

XX Allen SM, Allen MB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarzynski WC;

XX WPI; 2003-201509/19.

XX N-PSDB; ACC00801.

XX Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity; caleosin-like activity; useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

XX Claim 12; Page 369-371; 542pp; English.

XX The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
 CC activated protein (MAP)-kinase activity, LIPID-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity and KCC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to  
 CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention

XX Sequence 707 AA;

Query Match 29.9%; Score 1046.5; DB 6; Length 707;  
 Best Local Similarity 36.8%; Pred. No. 1.4e-87;

Matches 285; Conservative 88; Mismatches 182; Indels 219; Gaps 30;

QY 17 NMLGSLSPKMKKATLS-AATVPTTFMSP-----SQSHLSNFGMXY 55  
 DB 5 NLLGSLSPKMKKATLS-AATVPTTFMSP-----SQSHLSNFGMXY 55  
 QY 56 CYGVGNGNPH---SPLTVAPLKSD-----GSLCILEALKRSQTVWVPTSSPKLEDF 104  
 DB 64 SYGIYE-AFHNNNSINTQDWKENYNSQNLGTSQNNQNNHQQQQ---QPKLEN 115  
 QY 105 FLGGATMGTHGYGSHRGSLSDSIYNSQNAEQPNRDLSPFRQGGHMSVQTHPYYSGL 164  
 DB 116 FLGG-----HSFGEHE-----QF-----YGGN 131  
 QY 165 LACHGLYQAPLEBEET-----KETHVDCSSIMPQTEGLKXWVAPTRFSTHQVLEQ 217  
 DB 132 NSASTYMPAPQVLAGGGGGGSSNSNTSSISGLSM---IKTWL---RNPZPHSENNN 181  
 QY 218 QMNCGMGNERNGVSLGSGCGELQSLSLMSFGSSQSCCTAPSGTDSVAVDAKKRGHAKL 277

Wed Mar 17 08:18:54 2004

us-10-024-632-2.rag

Page 16

Db 182 ENIN-----NNESGNTIRSSVOQTLISLMSSTGSS--TSLPLLTASY-----DNRESS 229  
Qy 278 LGOKP-----VHRKSIDPFGQRTSOYRGVTRHRMTGRYEALHWDNSCK 321  
Db 230 SDKKQPTSAALDSTOGAIETAPRKSIDPFGORTSIRGVTRHRMTGRYEALHWDNSCR 289  
Qy 322 KEGGTRKRGROYLGGYMEKARAYDLALKTWGSPSTHINSIENYOQLEBKMCMRSRQ 381  
Db 290 REGGTRKRGROYLGGYDKEKARAYDLALKTWGTTTTTNPISHYEKELEBKMCMTRQ 349  
Qy 382 EYVAHLRKKSGFSGASIRGVTRRHOGHRCARIGRVAGNKDLYLGTFSTOBEAAY 441  
Db 350 EYVASLARKSSGFSRGASIRGVTRRHOGHRCARIGRVAGNKDLYLGTFSTOBEAAY 409  
Qy 442 DYALIKFRGANATNPDISRIYDERIMASSNLAGEARKK-----DNDRPKDIT 492  
Db 410 DYALIKFRGLSAVTNPFMSRYDVKSILESTTLPIGGAARLKDMEQVELSVNGHRADQV 469  
Qy 493 DYNKSVVTS-----VNNETVOQAQNNNNNENDSEMKI-----VLFNHP-----SQQ 534  
Db 470 DH--SILMSHLFGGINNNYAGGTAHNNWHNAHAFHQPOCTMHPYGGRIWCKQE 527  
Qy 535 QOANGNG-----SDOKIMNCNYRNSAF---SMALODIGIDSV-----570  
Db 528 QODNSDAPHLSYSDIHQLOLGNNGTNNFHTNGLHPMLSMDSASIDNSSSSNSVYDG 587  
Qy 571 --GSGQHML-----DESSKIGTHFSNT-----SSLVTSLSSSSREASPEK 608  
Db 588 YGGGGGYNNVMPWGTTAVVASDGDQNPNSNHGFGNEIKALGYESVYGSATDSYAHARN 647  
Qy 609 RGRBLFPMPMETKIVNFI-----GTSVTSMLPS--PTVQMRPSPALISLH 653  
Db 648 ----LYYLTQOQSSSVDTVKASAYDQGSACNWTWPTALPTAHPRSTYSMALCH 696

Search completed: March 9, 2004, 10:45:34  
Job time : 63 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using ew model

Run on: March 9, 2004, 10:44:24 ; Search time 23 Seconds

(without alignments)  
1488.175 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502

Sequence: 1 MKRINESNNTDDGNHNLG.....RSPALISHLVFASTWTD 663

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA.\*

1: /cgnt2\_6/prodata/2/1aa/5A\_COMB.pep.\*  
2: /cgnt2\_6/prodata/2/1aa/5B\_COMB.pep.\*  
3: /cgnt2\_6/prodata/2/1aa/6A\_COMB.pep.\*  
4: /cgnt2\_6/prodata/2/1aa/6B\_COMB.pep.\*  
5: /cgnt2\_6/prodata/2/1aa/PCITUS\_COMB.pep.\*  
6: /cgnt2\_6/prodata/2/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1328	37.9	US-09-227-421-2	Sequence 2, Appl1
2	1328	37.9	US-09-479-855-2	Sequence 2, Appl1
3	1307	37.3	US-09-227-421-5	Sequence 5, Appl1
4	1307	37.3	US-09-479-855-5	Sequence 5, Appl1
5	517	14.8	US-08-700-152A-4	Sequence 4, Appl1
6	385	11.0	US-08-912-272-9	Sequence 9, Appl1
7	385	11.0	US-09-026-039-9	Sequence 9, Appl1
8	346	9.9	US-08-912-272-10	Sequence 10, Appl1
9	346	9.9	US-09-026-039-10	Sequence 10, Appl1
10	197	5.6	US-08-700-152A-1	Sequence 1, Appl1
11	192.5	5.5	US-08-912-272-11	Sequence 11, Appl1
12	192.5	5.5	US-09-026-039-11	Sequence 11, Appl1
13	191.5	5.5	US-08-700-152A-2	Sequence 2, Appl1
14	191.5	5.5	US-09-026-039-5	Sequence 5, Appl1
15	191.5	5.5	US-09-026-039-5	Sequence 5, Appl1
16	188	5.4	US-08-912-272-4	Sequence 4, Appl1
17	188	5.4	US-09-026-039-4	Sequence 4, Appl1
18	156.5	4.5	US-09-202-161B-3	Sequence 3, Appl1
19	153	4.4	US-09-198-119C-55	Sequence 55, Appl1
20	148.5	4.2	US-09-198-119C-87	Sequence 87, Appl1
21	147	4.2	US-09-533-029-22	Sequence 22, Appl1
22	146	4.2	US-09-198-119C-71	Sequence 71, Appl1
23	145	4.1	US-09-198-119C-53	Sequence 53, Appl1
24	141.5	4.0	US-09-533-029-66	Sequence 66, Appl1
25	140.5	4.0	US-09-533-029-82	Sequence 82, Appl1
26	138.5	4.0	US-09-533-029-12	Sequence 12, Appl1
27	137.5	3.9	US-09-300-672-2	Sequence 2, Appl1

28	136	3.9	US-09-198-119C-17	Sequence 17, Appl1
29	136	3.9	US-09-198-119C-47	Sequence 47, Appl1
30	135.5	3.9	US-09-198-119C-83	Sequence 83, Appl1
31	135	3.9	US-09-202-161B-1	Sequence 1, Appl1
32	134.5	3.8	US-09-202-161B-27	Sequence 27, Appl1
33	134.5	3.8	US-09-198-119C-51	Sequence 51, Appl1
34	134	3.8	US-09-533-029-34	Sequence 34, Appl1
35	133.5	3.8	US-09-198-119C-77	Sequence 77, Appl1
36	133	3.8	US-09-198-119C-93	Sequence 93, Appl1
37	132.5	3.8	US-08-912-272-29	Sequence 29, Appl1
38	132.5	3.8	US-09-026-039-29	Sequence 29, Appl1
39	132.5	3.8	US-09-198-119C-61	Sequence 61, Appl1
40	132.5	3.8	US-09-198-119C-69	Sequence 69, Appl1
41	132.5	3.8	US-09-198-119C-63	Sequence 63, Appl1
42	132	3.8	US-09-198-119C-91	Sequence 91, Appl1
43	131.5	3.8	US-08-894-731-4	Sequence 4, Appl1
44	131	3.7	US-09-198-119C-59	Sequence 59, Appl1
45	130.5	3.7	US-09-198-119C-57	Sequence 57, Appl1

#### ALIGNMENTS

```

RESULT 1
US-09-227-421-2
Sequence 2, Application US/09227421
Patent No. 6559357
GENERAL INFORMATION:
APPLICANT: Fischer, Robert L.
APPLICANT: Mizukami, Yukiko
TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
FILE REFERENCE: 023070-090700PC
CURRENT APPLICATION NUMBER: US/09/227,421
PRIOR FILING DATE: 1999-01-08
PRIORITY APPLICATION NUMBER: US 09/227,421
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 555
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-227-421-2

Query Match      37.9% Score 1328 DB 4; Length 555;
Best Local Similarity 43.5% Pred. No. 3.4e-123;
Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

QY 1 MKRINESNNTDDGNHNLGFSLSPL-MKM-----EATSAATVPTTFVM 43
DB 1 MAFPCDNDNNHSTNLTILGFSLSNMKMGGRGGRALYSSSTSSAATSSSVPLV 60
QY 44 SPSQSHLSFMCVCVGENGNFHSPLTVMPKSDGSCITLALKRSGQVMVPTSPKLE 103
DB 61 G---DNTSNFVCVCGSNPNNGIYSHMSVMPLRSGSLCLNEALNRSSHNHODSSPKVE 117
QY 104 DFLGATGATGTEYSHENG--LSLDSIYNSQNAEAPNDLLSGP-----ROCGHVSVC 157
DB 116 DFFG---THNNNSHREANDLSLDSLFYNTT---EPNTTNPQEFSPFQNRH---- 166
QY 158 THPIYSGIACGLYQAPLEETTTKETHVSCSSIMPMQTEGLKMWAPVREFTHQOYLE 217
DB 167 -----EEET-RNYGND-----PFLTNG----- 182
QY 218 QMNCGMNENBNGVSLGSGGEL-QSLTSLMSPGSSQSCVTAPS----- 261
DB 193 -----GSFVGV-----YGEFGQSLSLMSBPGSSSCITSHHQNQNQNSQNH 229
QY 262 -----GTSVAVDKAKGH---AKLGQKQPVHRKSITPFGQFQSYGQVTR 304
DB 230 QQISEALVETSVGFEITTTAAAKKRGQEDVVVVGQKQIVHRKSIDTFQGFQSYGQVTR 289

```

305 HRMTGRYEALMDNSCKKEGQTRKGRQVYLLGGYDMEERKAAAYDIALKTYGSPSTHINFS 364  
Db 290 HRMTGRYEALMDNSCKKEGQTRKGRQVYLLGGYDMEERKAAAYDIALKTYGSPSTHINFS 349  
Qy 365 IENYVOUEEMKMGSRGROEVVAHLRKRKSGFSRGASITRGVTRHQRHQRGVARIGVAGNK 424  
Db 350 AENYQKEIEMKMGSRGROEVVAHLRKRKSGFSRGASITRGVTRHQRHQRGVARIGVAGNK 409  
Qy 425 DLYLGTFTGQEBAAAYDVAALIKFRGANAVTNFDISRYDVERIMASSNLGELARRKD 484  
Db 410 DLYLGTFTGQEBAAAYDVAALIKFRGTNAVTNFDITRYDVRIMASSNLGELARRKD 467  
Qy 485 NDPRKCIDYKNSVYTSVNNETTVQVQAGNNNNENDSEMKVLFPHNSQOQOQANGSDQ 544  
Db 468 -----NNSIV--VRNTE-----DQ 479  
Qy 545 KIMNGYRNBSAFSVALODLIGDSVSGQHMLDESSKIGTHFNTSSLYTSLSSREA 604  
Db 480 TALNA-----VVEGSKNEY 494  
Qy 605 SPEKRGPSL--LFMPMPMETKIV-NPIGTSVTSMIPSPVQMRPSPALISHLPPFASWT 661  
Db 495 STEPERLSFPALFALPQVQKMGFSNGMGNNMSPMTSPNMLK-TVALTLPPQMPVFAWA 553  
Qy 662 DT 663  
Db 554 DS 555

RESULT 2  
US-09-479-855-2  
; Sequence 2, Application US/09479855  
; Patent No. 6639128  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Robert L.  
; APPLICANT: Mizukami, Yukiko  
; TITLE OF INVENTION: The Regents of the University of California  
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
; FILE REFERENCE: 023070-090720US  
; CURRENT APPLICATION NUMBER: US/09/479,855  
; CURRENT FILING DATE: 2000-01-07  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 555  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; OTHER INFORMATION: AINTEGUMENTA (ANT)  
US-09-479-855-2

Query Match 37.9%; Score 1328; DB 4; Length 555;  
Best Local Similarity 43.5%; Pred. No. 3,4e-123;  
Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

Qy 1 MKRINESNNTDDGNHMLGFSISPH-MKM-----EATSAATPTTFM 43  
Db 1 MKSPCDNDNNHNTTLGLFSLSSNMKMGGRGREALYSSSTSSAATSSSVPLVV 60  
Qy 44 SPQSHLSNFGMCYGVGENGFHSPFLVWPLKSDGSLCILEALRSQTVWVPSPKLE 103  
Db 61 G---DNTSNFGVCGSNNGIYSHMSYMPLRSDGSLCMEALNRSSHNSHHODSSPKVE 117  
Qy 104 DFLGATWGTHERYSGHERG--LSIDSIYNSQNAEAPNRDLISQPT---RQCGHNSVQ 157  
Db 118 DFFG---THNNNTSHKRAMDLSIDSLFYNTTH---EENTNTNQEFFSPQCTNH--- 166  
Qy 158 THPYSGLAGHGLYQAPLEETETKETHVSDCSSLMPQWTEGLKNMVAPTREFTTHQVLE 217  
Db 167 -----EERT--RNYGND-----PSLTHG----- 182  
Qy 218 QQANCGMGNERNGVLSVSGCEL-QSLSLSSPSQSSCVTAPS----- 261

Db 183 -----GSFNVG-----YGEFOOSLSLSSPSQSSCITGSHHNOONONHQSQNH 229  
Qy 262 -----GTDVAADVAKKRGH---AKGQKQPVHRKSIDTFGQRTSQRYGTR 304  
Db 230 QQISEALVETSVGFETTTMAAKKRGQEDVYVQKQIVHRKSIDTFGQRTSQRYGTR 289  
Qy 305 HRMTGRYEALMDNSCKKEGQTRKGRQVYLLGGYDMEERKAAAYDIALKTYGSPSTHINFS 364  
Db 290 HRMTGRYEALMDNSCKKEGQTRKGRQVYLLGGYDMEERKAAAYDIALKTYGSPSTHINFS 349  
Qy 365 IENYVOUEEMKMGSRGROEVVAHLRKRKSGFSRGASITRGVTRHQRHQRGVARIGVAGNK 424  
Db 350 AENYQKEIEMKMGSRGROEVVAHLRKRKSGFSRGASITRGVTRHQRHQRGVARIGVAGNK 409  
Qy 425 DLYLGTFTGQEBAAAYDVAALIKFRGANAVTNFDISRYDVERIMASSNLGELARRKD 484  
Db 410 DLYLGTFTGQEBAAAYDVAALIKFRGTNAVTNFDITRYDVRIMASSNLGELARRKD 467  
Qy 485 NDPRKCIDYKNSVYTSVNNETTVQVQAGNNNNENDSEMKVLFPHNSQOQOQANGSDQ 544  
Db 468 -----NNSIV--VRNTE-----DQ 479  
Qy 545 KIMNGYRNBSAFSVALODLIGDSVSGQHMLDESSKIGTHFNTSSLYTSLSSREA 604  
Db 480 TALNA-----VVEGSKNEY 494  
Qy 605 SPEKRGPSL--LFMPMPMETKIV-NPIGTSVTSMIPSPVQMRPSPALISHLPPFASWT 661  
Db 495 STEPERLSFPALFALPQVQKMGFSNGMGNNMSPMTSPNMLK-TVALTLPPQMPVFAWA 553  
Qy 662 DT 663  
Db 554 DS 555

RESULT 3  
US-09-227-421-5  
; Sequence 5, Application US/09227421  
; Patent No. 6559357  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Robert L.  
; APPLICANT: Mizukami, Yukiko  
; TITLE OF INVENTION: The Regents of the University of California  
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
; FILE REFERENCE: 023070-090700PC  
; CURRENT APPLICATION NUMBER: US/09/227,421  
; CURRENT FILING DATE: 1999-01-08  
; PRIOR APPLICATION NUMBER: US 09/227,421  
; PRIOR FILING DATE: 1999-01-08  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 548  
; TYPE: PRT  
; ORGANISM: Brassica napus  
US-09-227-421-5

Query Match 37.3%; Score 1307; DB 4; Length 548;  
Best Local Similarity 43.6%; Pred. No. 4.1e-121;  
Matches 309; Conservative 77; Mismatches 103; Indels 220; Gaps 25;

Qy 9 NTDGNNHMLGFSISPH-----KVENSAATVPTTFMSPSQ-----SHLSNFGMCY 57  
Db 6 DNDSDVTTTLGLFSLSSNMKMGGRGREALYSSSSSVATSS--SVPLVVGDNSSNYGVCY 63  
Qy 58 GVG-ENGNFHSPFLVWPLKSDGSLCILEALRS-----QTVWVPSPLEDFLGA 109  
Db 64 GSNLAARENYQMSYMPLRSDGSLCMEALNRSSHNSHHNSQV-----SSPMEGF----- 115  
Qy 110 TWGTHRYG-SHERG--LSIDSIYNSQNAEAPNRDLISQPRQGHNSVQTHPYSGLA 166



Db 116 -FGTHNHTSHKEMDLSLDFYNTTHA---PNNNTNFOEF-----FSPQTRNNH----- 163  
 Qy 167 CHGLYQALBEEETKETHTVSDCSLMPQMTGKKNVAFPRFSTHQVLEQMGCMGN 226  
 Db 164 -----EBET--RYEND-----PGLTHG----- 179  
 Qy 227 ERNGVSLGVSOGGEL--QSLISMSPGSQSCVTA-----P 260  
 Db 180 ---GGSFNVGYYGEFQOSLSISMSPGSQSCITASHHQNQTONHOQISEALVETSGAFE 236  
 Qy 261 SGTDSVAVDAKRG--HAKLGOKOPYHRKSIDTFQORTSOYRGVTRHRTGRYAHLMDS 319  
 Db 237 TTTMAAAAKKRGQGEVVGQKQIVHRKSIDTFQORTSOYRGVTRHRTGRYAHLMDS 296  
 Qy 320 CKKEGQTRKGRQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 379  
 Db 297 FKKEGHRKRGQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 356  
 Qy 380 ROEYVAHLRRKSGFSRGSASIRGVTRHQRWQARIGVAGNKDLYLGFTQGEAAE 439  
 Db 357 ROEYVAHLRRKTSRGSASIRGVTRHQRWQARIGVAGNKDLYLGFTQGEAAE 416  
 Qy 440 AYDVAALKRFGANAVTNFDISRVERIMASNLGELARRKKNDPRKIDYKSAV 499  
 Db 417 AYDVAALKRFGTNAVTFDITRYVDRIKMASNTLLSGEMARRNS-----NSIV 465  
 Qy 500 TSVNNE---TVQVQAGNNNNENDESMKVLFNHPSQQQOANGSDOKIMNCGYRNSA 556  
 Db 466 RNISDEEALTAIVVAGSNKEV-----GSPERVLSF----- 496  
 Qy 557 FSMALQDLIGDVSQGHNMLEDSSKIGTHFNTSLSVTSLSRSSRASPBRGSLIFP 616  
 Db 497 -----PTIFA 501  
 Qy 617 MPMEETKI--VNPIGTSVSWLPSPPTVQMRPSAISLSHPVFAWTD 663  
 Db 502 LPQVGPXMGANVVG--NMSWTTNPNADLK--TVSLTLQPMVFAWADS 548

RESULT 4  
 US-09-479-855-5  
 ; Sequence 5, Application US/09479855  
 ; Patent No. 6639128  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fischer, Robert L.  
 ; APPLICANT: Mizukami, Yukiko  
 ; APPLICANT: The Regents of the University of California  
 ; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
 ; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants  
 ; FILE REFERENCE: 023070-090720US  
 ; CURRENT APPLICATION NUMBER: US/09/479,855  
 ; CURRENT FILING DATE: 2000-01-07  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 548  
 ; TYPE: PRT  
 ; ORGANISM: Brassica napus  
 ; OTHER INFORMATION: canola AINTEGUMENTA (ANT)  
 US-09-479-855-5

Query Match 37.3%; Score 1307; DB 4; Length 548;  
 Best Local Similarity 43.6%; Pred. No. 4, 1e-121;  
 Matches 309; Conservative 77; Mismatches 103; Indels 220; Gaps 25;

Qy 110 TWGTHEYG-SHERG--ISLDSIYNSQNAEQPNRDLISQFROQGHNSVQTHPYSGLA 166  
 Db 116 -FGTHNHTSHKEMDLSLDFYNTTHA---PNNNTNFOEF-----FSPQTRNNH----- 163  
 Qy 167 CHGLYQALBEEETKETHTVSDCSLMPQMTGKKNVAFPRFSTHQVLEQMGCMGN 226  
 Db 164 -----EBET--RYEND-----PGLTHG----- 179  
 Qy 227 ERNGVSLGVSOGGEL--QSLISMSPGSQSCVTA-----P 260  
 Db 180 ---GGSFNVGYYGEFQOSLSISMSPGSQSCITASHHQNQTONHOQISEALVETSGAFE 236  
 Qy 261 SGTDSVAVDAKRG--HAKLGOKOPYHRKSIDTFQORTSOYRGVTRHRTGRYAHLMDS 319  
 Db 237 TTTMAAAAKKRGQGEVVGQKQIVHRKSIDTFQORTSOYRGVTRHRTGRYAHLMDS 296  
 Qy 320 CKKEGQTRKGRQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 379  
 Db 297 FKKEGHRKRGQVYLGVDMEKARAYDLAALKYWGSPSTHINSIENYQVLEBKMS 356  
 Qy 380 ROEYVAHLRRKSGFSRGSASIRGVTRHQRWQARIGVAGNKDLYLGFTQGEAAE 439  
 Db 357 ROEYVAHLRRKTSRGSASIRGVTRHQRWQARIGVAGNKDLYLGFTQGEAAE 416  
 Qy 440 AYDVAALKRFGANAVTNFDISRVERIMASNLGELARRKKNDPRKIDYKSAV 499  
 Db 417 AYDVAALKRFGTNAVTFDITRYVDRIKMASNTLLSGEMARRNS-----NSIV 465  
 Qy 500 TSVNNE---TVQVQAGNNNNENDESMKVLFNHPSQQQOANGSDOKIMNCGYRNSA 556  
 Db 466 RNISDEEALTAIVVAGSNKEV-----GSPERVLSF----- 496  
 Qy 557 FSMALQDLIGDVSQGHNMLEDSSKIGTHFNTSLSVTSLSRSSRASPBRGSLIFP 616  
 Db 497 -----PTIFA 501  
 Qy 617 MPMEETKI--VNPIGTSVSWLPSPPTVQMRPSAISLSHPVFAWTD 663  
 Db 502 LPQVGPXMGANVVG--NMSWTTNPNADLK--TVSLTLQPMVFAWADS 548

RESULT 5  
 US-08-700-152A-4  
 ; Sequence 4, Application US/08700152A  
 ; Patent No. 5994622  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jofuku, K. Diane  
 ; APPLICANT: Okamoto, Jack K.  
 ; TITLE OF INVENTION: Methods for Improving Seeds  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSER: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, Eighth Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/700,152A  
 ; FILING DATE: 20-AUG-1996  
 ; CLASSIFICATION: 800  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bastian, Kevin L.  
 ; REGISTRATION NUMBER: 34,774  
 ; REFERENCE/DOCKET NUMBER: 023070-067200US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 576-0200  
 ; TELEFAX: (415) 576-0300

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

RESULT 7  
 US-09-026-039-9  
 Sequence 9, Application US/09026039  
 Patent No. 6329667  
 GENERAL INFORMATION:  
 APPLICANT: Jotiku, K. Diane  
 APPLICANT: Okamoto, Jack K.  
 TITLE OF INVENTION: Methods for Improving Seeds  
 NUMBER OF SEQUENCES: 103  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/026,039  
 FILING DATE: 19-FEB-1998  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/912,272  
 FILING DATE: 15-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA: US 08/700,152  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 77 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
MOLECULE TYPE: linear  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..77  
OTHER INFORMATION: /note= "ANT-R1 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 44..59  
OTHER INFORMATION: /note= "putative ANT-R1 amphipathic  
OTHER INFORMATION: alpha-helix (SEQ ID NO:37)"  
US-09-026-039-9

Query Match 11.0%; Score 385; DB 4; Length 77;  
Best Local Similarity 89.6%; Pred. No. 1e-30; 5; Indels 0; Gaps 0;  
Matches 69; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 296 TSOYRGVTRHMTGRYBAHLMDNSCKEKGQTRKGRQVYLGGMEEKARAYDLALKTW 355  
DB 1 TSOYRGVTRHMTGRYBAHLMDNSFKKEGHSRKGRQVYLGGMEEKARAYDLALKTW 60

QY 356 GPSTHTNFSANVQKL 372  
DB 61 GPSTHTNFSANVQKEI 77

RESULT 8  
US-08-912-272-10  
Sequence 10, Application US/08912272  
Patent No. 6093874  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamura, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,272  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067220US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
MOLECULE TYPE: peptide  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..69  
OTHER INFORMATION: /note= "ANT-R2 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 37..51  
OTHER INFORMATION: /note= "putative ANT-R2 amphipathic  
OTHER INFORMATION: alpha-helix"  
US-08-912-272-10

Query Match 9.9%; Score 346; DB 3; Length 69;  
Best Local Similarity 94.2%; Pred. No. 6.6e-27; 2; Indels 0; Gaps 0;  
Matches 65; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 398 ASYRGVTRHHQSGQARIGRVAGNKDLYGTFTSTGEAAEAYDVAIKRGANAYTNF 457  
DB 1 ASYRGVTRHHQSGQARIGRVAGNKDLYGTFTSTGEAAEAYDVAIKRGANAYTNF 60

QY 458 DISRYDVER 466  
DB 61 DISRYDVDR 69

RESULT 9  
US-09-026-039-10  
Sequence 10, Application US/09026039  
Patent No. 6329567  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamura, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,039  
FILING DATE: 19-FEB-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,272  
FILING DATE: 15-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152

FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..69  
OTHER INFORMATION: /note= "ANT-R2 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 37..51  
OTHER INFORMATION: /note= "putative ANT-R2 amphipathic  
OTHER INFORMATION: alpha-helix"  
US-09-026-039-10

Query Match 9.9%; Score 346; DB 4; Length 69;  
Best Local Similarity 94.2%; Pred. No. 6.6e-27;  
Matches 65; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 358 ASIVGVTRHHQHGHWARIGVAGNKLVLGTSTGEAAEAYDAIKRGANAYTNF 457  
DB 1 ASIVGVTRHHQHGHWARIGVAGNKLVLGTSTGEAAEAYDAIKRGANAYTNF 60

QY 458 DISRYDVR 466  
DB 61 DITRYDVR 69

RESULT 10  
US-08-700-152A-1  
Sequence 1, Application US/08700152A  
Patent No. 5994622  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamuro, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,152A  
FILING DATE: 20-AUG-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067200US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0300  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 67 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..67  
OTHER INFORMATION: /note= "APETALA 2 (AP2) of Arabidopsis  
OTHER INFORMATION: direct repeat domain AP2-R1 consisting  
OTHER INFORMATION: of amino acids 129-195 of the AP2  
OTHER INFORMATION: protein"  
US-08-700-152A-1

Query Match 5.6%; Score 197; DB 2; Length 67;  
Best Local Similarity 51.9%; Pred. No. 4.5e-12;  
Matches 40; Conservative 12; Mismatches 15; Indels 10; Gaps 2;

QY 296 TSQYRGVTRHRTGTYEALHMDNSCKEGQTRKGRQVTLGYDMEKAAAYDLAALKY 355  
DB 1 SSQYRGVTFYRTGRTGWESHWD-C-----GQVYLGFTDAHAARAAYDRAAIKFR 50

QY 356 GPSTHINPSIENYQVL 372  
DB 51 GVEADINFINIDYDDL 67

RESULT 11  
US-08-912-272-11  
Sequence 11, Application US/08912272

Patent No. 6033874  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamuro, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,272  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 800

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997

PRIOR APPLICATION NUMBER: US 08/700,152  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067220US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 67 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:

NAME/KEY: Peptide  
LOCATION: 1..67  
OTHER INFORMATION: /note= "RAP2.7-R1 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..49  
OTHER INFORMATION: /note= "putative RAP2.7-R1 amphipathic"  
OTHER INFORMATION: alpha helix (SEQ ID NO:36)"  
US-08-912-272-11

Query Match 5.5%; Score 192.5; DB 3; Length 67;  
Best Local Similarity 50.0%; Pred. No. 1.3e-11;  
Matches 40; Conservative 14; Mismatches 13; Indels 13; Gaps 3;

QY 296 TSGYRGVTRHWTGRYEAHLMDNSCKKEGQTRKRGQVYLGSDMEKAAAYDLAALKYW 355  
DB 1 SSGYRGVTRHWTGRYEAHLMDNSCKKEGQTRKRGQVYLGSDMEKAAAYDLAALKYW 50  
QY 356 GPSTHINFSIENYQVLEEM 375  
DB 51 GVDADINFTLGDYE---EDM 67

RESULT 12  
US-09-026-039-11

Sequence 11, Application US/09026039  
Patent No. 6329567  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,039  
FILING DATE: 19-FEB-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,272  
FILING DATE: 15-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 67 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide

LOCATION: 1..67  
OTHER INFORMATION: /note= "RAP2.7-R1 direct repeat"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..49  
OTHER INFORMATION: /note= "putative RAP2.7-R1 amphipathic"  
OTHER INFORMATION: alpha helix (SEQ ID NO:36)"  
US-09-026-039-11

Query Match 5.5%; Score 192.5; DB 4; Length 67;  
Best Local Similarity 50.0%; Pred. No. 1.3e-11;  
Matches 40; Conservative 14; Mismatches 13; Indels 13; Gaps 3;

QY 296 TSGYRGVTRHWTGRYEAHLMDNSCKKEGQTRKRGQVYLGSDMEKAAAYDLAALKYW 355  
DB 1 SSGYRGVTRHWTGRYEAHLMDNSCKKEGQTRKRGQVYLGSDMEKAAAYDLAALKYW 50  
QY 356 GPSTHINFSIENYQVLEEM 375  
DB 51 GVDADINFTLGDYE---EDM 67

RESULT 13  
US-08-700-152A-2

Sequence 2, Application US/08700152A  
Patent No. 5994692  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,152A  
FILING DATE: 20-AUG-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067200US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Domain  
LOCATION: 1..68  
OTHER INFORMATION: /note= "APETALA 2 (AP2) of Arabidopsis"  
OTHER INFORMATION: direct repeat domain AP2-R2 consisting  
of amino acids 221-288 of the AP2  
OTHER INFORMATION: protein"  
US-08-700-152A-2

Query Match 5.5%; Score 191.5; DB 2; Length 68;  
Best Local Similarity 60.3%; Pred. No. 1.6e-11;  
Matches 41; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 398 ASIVRGVTRHHQHGWRARIGRVAGNKDLYLGTFTSQEBAEAYVAIKRGANAVTNF 457  
1 SSKYRGVTL-HKCGRWEARMGQFLGKKYVYLGFDTEVEAARAVDAIKKNGKDAVTNF 59  
DB 458 DISRYDVE 465  
60 DPSIYDEE 67

RESULT 14  
US-08-912-272-5  
Sequence 5, Application US/08912272  
Patent No. 6038874  
GENERAL INFORMATION:  
APPLICANT: Jotuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,272  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067220US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..68  
OTHER INFORMATION: /note="AP2-R2 direct repeat at  
OTHER INFORMATION: positions 221 to 288"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..50  
OTHER INFORMATION: /note="putative AP2-R2 amphipathic  
OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"  
US-08-912-272-5

Query Match 5.5%; Score 191.5; DB 3; Length 68;  
Best Local Similarity 60.3%; Pred. No. 1.6e-11;  
Matches 41; Conservative 8; Mismatches 18; Indels 1; Gaps 1;

DB 1 SSKYRGVTL-HKCGRWEARMGQFLGKKYVYLGFDTEVEAARAVDAIKKNGKDAVTNF 59  
QY 458 DISRYDVE 465  
DB 60 DPSIYDEE 67

RESULT 15  
US-09-026-039-5  
Sequence 5, Application US/09026039  
Patent No. 6329567  
GENERAL INFORMATION:  
APPLICANT: Jotuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/026,039  
FILING DATE: 19-FEB-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,272  
FILING DATE: 15-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/879,827  
FILING DATE: 20-JUN-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/700,152  
FILING DATE: 20-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-067230US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..68  
OTHER INFORMATION: /note="AP2-R2 direct repeat at  
OTHER INFORMATION: positions 221 to 288"  
FEATURE:  
NAME/KEY: Region  
LOCATION: 33..50  
OTHER INFORMATION: /note="putative AP2-R2 amphipathic  
OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"  
US-09-026-039-5

Query Match 5.5%; Score 191.5; DB 4; Length 68;  
Best Local Similarity 60.3%; Pred. No. 1.6e-11;  
Matches 41; Conservative 8; Mismatches 18; Indels 1; Gaps 1;

Wed Mar 17 08:18:55 2004

us-10-024-632-2.rat

Page 9

Db 1 SSKYRGVTL-HKCGWEARNQGLGKKVYIUGLFDTEVEARARAYDKAIAIKONGDAVTNF 59

QY 458 DISRYDVE 465

Db 60 DPSIYDEE 67

Search completed: March 9, 2004, 10:48:10  
Job time : 25 secs

Wed Mar 17 08:18:55 2004

us-10-024-632-2.rnt

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 13, 2004, 00:46:03 ; Search time 98 Seconds  
(without alignments)  
3754.413 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502

Sequence: 1 MKRINBSNNTDDGNHNLG.....RSPALSLHLPVPSWTDT 663

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model DEV=xlh  
-Q=/cgn2\_1/USPTO/US10024632/runat\_09032004\_101208.10797/app query.fasta\_1.839  
-DB=Issued\_Patents\_NA -QEMT=fastap -SUFFIX=rnt -MINMATCH=0.1 -LOOFCU=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=numan40.cdt  
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10024632@cgn2\_1.1.44@runat\_09032004\_101208.10797 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEUSER -NEG SCORES=0 -WART -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MARK\_TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6  
-YGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1:	/cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2:	/cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3:	/cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4:	/cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5:	/cgn2_6/ptodata/2/ina/6CTUS.COMB.seq:*
6:	/cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1338	37.9	2148	4	US-09-227-421-1
2	1358	37.9	2148	4	US-09-479-855-1
3	1307	37.3	1738	4	US-09-227-421-4
4	1307	37.3	1738	4	US-09-479-855-4
5	517	14.8	1680	2	US-08-700-152A-3
6	292	8.3	11721	4	US-09-026-039-3
7	292	8.3	11721	4	US-09-026-039-3
8	263	7.5	101	4	US-09-113-294A-6727
9	158.5	4.5	888	4	US-09-198-119C-54
10	156.5	4.5	1405	4	US-09-202-161B-6
11	154.5	4.4	1669	4	US-09-026-039-1
12	148.5	4.2	854	4	US-09-198-119C-86

13	148	4.2	950	4	US-09-198-119C-70	Sequence 70, Appl
14	147	4.2	874	4	US-09-198-119C-52	Sequence 52, Appl
15	147	4.2	1055	4	US-09-533-029-21	Sequence 21, Appl
16	145.5	4.2	1577	4	US-09-533-029-11	Sequence 11, Appl
17	142	4.0	294	4	US-09-313-294A-6087	Sequence 6087, Ap
18	141.5	4.0	1281	4	US-09-533-029-65	Sequence 65, Appl
19	140.5	4.0	751	4	US-09-533-029-81	Sequence 81, Appl
20	140	4.0	1500	3	US-09-300-672-1	Sequence 1, Appl
21	139.5	4.0	832	4	US-09-198-119C-82	Sequence 82, Appl
22	139	4.0	8146	4	US-09-975-594-725	Sequence 725, App
23	139	4.0	11917	4	US-09-566-921-32	Sequence 32, Appl
24	137.5	3.9	913	4	US-09-533-029-33	Sequence 33, Appl
25	137	3.9	1499	3	US-09-300-672-3	Sequence 3, Appl
26	136	3.9	632	4	US-09-198-119C-18	Sequence 18, Appl
27	136	3.9	632	4	US-09-198-119C-46	Sequence 46, Appl
28	135	3.9	933	4	US-09-202-161B-4	Sequence 4, Appl
29	134.5	3.8	174	4	US-09-202-161B-21	Sequence 21, Appl
30	134.5	3.8	682	4	US-09-198-119C-92	Sequence 92, Appl
31	134.5	3.8	884	4	US-09-198-119C-50	Sequence 50, Appl
32	133.5	3.8	830	4	US-09-198-119C-84	Sequence 84, Appl
33	133.5	3.8	884	4	US-09-198-119C-76	Sequence 76, Appl
34	132.5	3.8	887	4	US-09-198-119C-68	Sequence 68, Appl
35	132.5	3.8	889	4	US-09-198-119C-62	Sequence 62, Appl
36	132.5	3.8	953	4	US-09-198-119C-60	Sequence 60, Appl
37	132	3.8	793	4	US-09-198-119C-80	Sequence 80, Appl
38	131.5	3.8	1132	3	US-08-894-731-3	Sequence 3, Appl
39	131	3.7	769	4	US-09-198-119C-58	Sequence 58, Appl
40	130.5	3.7	755	4	US-09-198-119C-80	Sequence 80, Appl
41	130.5	3.7	806	4	US-09-198-119C-78	Sequence 78, Appl
42	130.5	3.7	876	4	US-09-198-119C-86	Sequence 86, Appl
43	130.5	3.7	1132	4	US-09-198-119C-56	Sequence 56, Appl
44	130	3.7	937	4	US-09-301-666A-5	Sequence 5, Appl
45	130	3.7	937	4	US-09-301-217-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-227-421-1  
Sequence 1, Application US/09227421  
Patent No. 6559357  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
FILE REFERENCE: 023070-090700PC  
CURRENT APPLICATION NUMBER: US/09/227,421  
CURRENT FILING DATE: 1999-01-08  
PRIOR APPLICATION NUMBER: US-09/727,421  
PRIOR FILING DATE: 1999-01-08  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 2148  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (269)..(1936)  
OTHER INFORMATION: AINTEGUMENTA (ANT)  
NAME/KEY: misc feature  
LOCATION: (1105)..(1339)  
OTHER INFORMATION: encodes first AP2 domain  
NAME/KEY: misc feature  
LOCATION: (1340)..(1414)  
OTHER INFORMATION: encodes linker region  
NAME/KEY: misc feature  
LOCATION: (1415)..(1621)  
OTHER INFORMATION: encodes second AP2 domain  
US-09-227-421-1





QY 624 ileval---AsnProIleGlyThrSerValThrSerIleProSerProThrValGln 642  
 DB 1814 ATGTTCGACCAATATGCGGAAATATGATGCTTGGACATCAACCTATATGCTGAG 1873  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerThrPhe 662  
 DB 1874 CTTAAG---ACCGTCCTTACTTTCCTCAGATGCCGGTTTTCGCTGCTGCTGAT 1930  
 QY 663 Thr 663  
 DB 1931 TCT 1933  
 RESULT 2  
 US-09-479-855-1  
 : Sequence 1, Application US/09479855  
 : Patent No. 6639128  
 : GENERAL INFORMATION:  
 : APPLICANT: Fischer, Robert L.  
 : APPLICANT: Mizukami, Yukiko  
 : TITLE OF INVENTION: The Regents of the University of California  
 : TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility  
 : FILE REFERENCE: 023070-090720US  
 : CURRENT APPLICATION NUMBER: US/09/479,855  
 : NUMBER OF FILING DATE: 2000-01-07  
 : NUMBER OF SEQ ID NOS: 8  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 1  
 : LENGTH: 2148  
 : TYPE: DNA  
 : ORGANISM: Arabidopsis thaliana  
 : FEATURE: INFORMATION: AINTEGUMENTA (ANT) cDNA  
 : NAME/KEY: CDS  
 : LOCATION: (269)..(1936)  
 : OTHER INFORMATION: AINTEGUMENTA (ANT)  
 US-09-479-855-1  
 Alignment Scores:  
 Pred. No.: 3,44e-130 Length: 2148  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.32% Indels: 224  
 DB: 4 Gaps: 21  
 US-10-024-632-2 (1-663) x US-09-479-855-1 (1-2148)  
 QY 1 MetIyAaGileAaNgIeAaNsIaNsThraSphApGlyAaNsAaNsIaNsThraPleuGly 20  
 DB 269 ATGAAGTCTTTTGGATATGATGATATATATATCATAGCAACAGACTAATTTGTTAGG 328  
 QY 21 PheSerLeuSerProHis---MetIyMet----- 29  
 DB 329 TTCTCATGCTTTCATATGATATGATATGATATGATATGATATGATATGATATGATATG 388  
 QY 30 -----GluAlaThrSerAlaAlaThrValProThrThrPheIyMet 43  
 DB 369 TCATCTCAACTTCTTCAGCTGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 448  
 QY 44 SerProSerGlnSerHisLeuSerAaNsPheGlyMetCysIyGlyValGlyGluAaNgIy 63  
 DB 449 GGT-----GACAACTAGCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 499  
 QY 64 AaNsPheAaSerProLeuThrValMetProLeuIySerAaPglySerLeuCysIleleu 83  
 DB 500 GGAATCTATTCACATGCTGTGATGCCACATGATGATGATGATGATGATGATGATGATG 559  
 QY 84 GluAlaLeuIyAaNgIeAaNsIaNsThraSphApGlyAaNsAaNsIaNsThraPleuGly 103  
 DB 560 GAAGCTCTCAACAGATCTTCTCACTGGAATCAACATCAAGATTCATCTCTCAAGGTGAG 619  
 QY 104 AspPheLeuGlyIyAlaThrMetGlyThrHisGlyIyGlySerHisGluAaNgIy--- 122

DB 620 GATTTCTTTGG3-----ACCATACAAACAACACATCTCACAAAGAACCATG 667  
 QY 123 ---LeuSerLeuAaPserIleIyTyraSnsGlnAaGluAlaGlnProAaNg 141  
 DB 668 GATCTTACTTATGATGATTTATCTTACAAACACATCAT-----GAGCCCAACAG 718  
 QY 142 AspleuLeuSerGlnProPhe-----ArgGlnGlnIyHisMetSerValGln 157  
 DB 719 ACTACAACTTTCAGAGCTTTAGCTTCCCTCAACACCAAAACCAT----- 766  
 QY 158 ThrHisProIyTyraSerGlyLeuAlaCysHisGlyLeuTyraGlnAlaProLeuGlu 177  
 DB 767 -----GAGGAA 772  
 QY 178 GluThrThrIySgluThrHisValSerAaPglySerSerLeuMetProGlnMetThrGlu 197  
 DB 773 GAACCT-----AGAAATTCAGGAATGAC-----CCTAGTTTGACACAT 811  
 QY 198 GlyLeuIyAaNsThrValAlaProThrArgGluPheSerThrHisGlnGlnValleuGlu 217  
 DB 812 GGA----- 814  
 QY 218 GlnGlnMetAaNsCysGlyMetCysGluAaNsGluAaNgIyValSerLeuGlySerValGly 237  
 DB 815 GGGCTTTTAAATGTAGGGGATATAGGGAATTTCA----- 850  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 DB 851 -----CAGTCACTGAGCTTATTCATGAGCCTGGGTCAATCTTACGTGCATC 898  
 QY 258 ThrAlaProSer----- 261  
 DB 899 ACTGGCTCACACACACCAACCAAAACCAAAACCAACCAACCAACCAACCAACCAACCA 958  
 QY 262 -----GlyThrAaPserValAlaVal 268  
 DB 959 CAGATCTCTGAAGCTCTTGTGAGACAAAGCTGGGTTTGAGACACACAAATGGCGCT 1018  
 QY 269 AspaAlaIyLeuAaNgIyHis-----AlaIyLeuGlyGlnIyGlnProValHis 285  
 DB 1019 GCGAAGAAAGAAAGGGGCAAGAGATGTGATGTTGGTCAGAAACAGATTTGTCAT 1078  
 QY 286 ArgIySerIleAaPThrPheGlyGlnArgThrSerGlnTyraArgIyValThrArgHis 305  
 DB 1079 AGAAATCTATGATGATCTTTGGACACAGACTTCTCAATACCGAGGCGTTTCAAGACAT 1138  
 QY 306 ArgThrThrGlyAaGlyGluAlaHisLeuThrAaPserSerCysIySgluGlyGln 325  
 DB 1139 AGATGACCTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1198  
 QY 326 ThrArgIyGlyAaGlnValIyIleuGlyGlyTyraSerMetGlnGlnIyAlaIaArg 345  
 DB 1199 AGTAAAGAAAGAAAGCAAGTTTATCTGGAGGTTTGAATGAGAGAAAGCTGCTGCA 1258  
 QY 346 AlaTyraPleuAlaAlaLeuIyTyraGlyProSerThrHisIleAaPserIle 365  
 DB 1259 GCATATGATCTTCTGCACTCAAGTATCGGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1318  
 QY 366 GluAaNsTyraGlnValGlnleuGluMetIyAaNsMetSerArgGlnGlyTyraAla 385  
 DB 1319 GAGAAATTCAGAAAGAAAGTTGAGACATGAAAGAAACATGATGACAAATAATGTTGA 1378  
 QY 386 HisLeuAaGlyAaSerSerGlyPheSerArgIyAlaSerIleTyraArgIyValThr 405  
 DB 1379 CATTTAGAAAGAAAGCAAGTGTCTTCTAGGGGCTCTCTCTCTCTCTCTCTCTCTCTCT 1438  
 QY 406 ArgHisHisGlnHisGlyAaGlyGlnAlaArgIleGlyAaValAlaGlyAaNsIyAaP 425  
 DB 1439 AGACATCACAGATGAGAGTGGCAAGCAAGATGTTGATGATGATGATGATGATGATGAT 1498  
 QY 426 LeuTyraLeuGlyThrPheSerThrGlnGlnGluAlaAlaGluAlaTyraPValAlaAla 445

```

Db 1499 CTCTACCTTGGACTTTGGACCCAGAAAGAGCTGACAGACTTTCATGATGACAGCA 1558
QY 446 11lelySphearGlyAlaAsnAlaValThrAsnPhenAspIleSerArgTyrAspValGlu 465
1559 ATTAAAGTTCCGTGACAAATAGCTGATCTACTATCATCAGAGGTGATGATTTGAT 1618
QY 466 ArgIleMetAlaSerSerAsnLeuAlaGlyGluLeuAlaArgArgTyrSlyAspAsn 485
1619 CGTATCATGCTGTATGACACACTTCTGTCTGAGAGTTACGCGAAGAAC----- 1669
QY 486 AspProArgAsnLysAspIleAspTyrAsnLysSerValValThrSerValAsnAsnGlu 505
1670 -----AACACAGCATTTGTC-----GTCAGAAATACAT 1696
QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGluTyrPlySme 525
1697 GAA----- 1699
QY 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545
1700 -----GACCAAAAC 1708
QY 546 11leuetaCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565
1709 GCTCTAAATGCT----- 1720
QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerLysIleGly 585
1720 ----- 1720
QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluAlaSer 605
1721 -----GTGTGGAAGGTGCTTCACAAAGAGTCAGT 1753
QY 606 ProGluLysArgGlyProSerLeu-----LeuPheProMetProPrometGluThrLys 623
1754 ACTCCCGAGAGACTTGGAGTTTCCGCGCGATTTTCGCGCTTCAAGTTAATCAAAG 1813
QY 624 11leval-----AsnProIleGlyThrSerValThrSerTyrLeuProSerProThrValGln 642
1814 ATGTTCCGATCAATATGCGCGAATATAGCTCTTCAATCAAAACCTTAATGCTGAG 1873
QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrPThrAsp 662
1874 CTAAAG---ACCGTGGCTCTTAATTTGCTCAGATGCCGTTTTCGCTGCTGGCTGAT 1930
QY 663 Thr 663
1931 TCT 1933

```

## RESULT 3

US-09-227-421-4

Sequence 4, Application US/09227421

Patent No. 6559357

GENERAL INFORMATION:

APPLICANT: Fischer, Robert L.

APPLICANT: Mizukami, Yukiko

TITLE OF INVENTION: The Regents of the University of California

TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility

FILE REFERENCE: 023070-090700PC

CURRENT FILING DATE: 1999-01-08

PRIOR APPLICATION NUMBER: US 09/227,421

PRIOR FILING DATE: 1999-01-08

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 4

LENGTH: 1738

TYPE: DNA

ORGANISM: Brassica napus

FEATURE:

NAME/KEY: CDS

```

; LOCATION: (1)-(1647)
; OTHER INFORMATION: BANT, ANT ortholog from Brassica napus (Canola)
US-09-227-421-4
Alignment Scores:
Pred. No.: 4,01e-128 Length: 1738
Score: 1307.00 Matches: 308
Percent Similarity: 54.52% Conservative: 78
Best Local Similarity: 43.50% Mismatches: 104
Query Match: 37.32% Indels: 218
DB: 4 Gaps: 25
US-10-024-632-2 (1-663) x US-09-227-421-4 (1-1738)
QY 9 AsnThrAspAspGlyAsnAsnHisAsnTyrLeuGlyPheSerLeuSerProHisMet--- 27
16 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 75
QY 28 -----LysMetGluAlaThrSerAlaAlaThrValProThrThrPhe 41
76 AAATGGGTGGTGGAGAGCTCTTACTCATCTTCGTCTTCACTTCACTTCTCTCTCTCTCT 132
QY 42 TyrMetSerProSerGln-----SerHisLeuSerAsnProMetGlyMetCysTyr 57
133 ---TGTGTCCACACACACTTGTGTGTGGCACAACACTGACCACTATGAGTTGCTAC 189
QY 58 GlyValGly---GluAsnGlyAsnPhenHisSerProLeuThrValMetProLeuLysSer 76
190 GGTTCATTAATGACGATGAGGAATGATTCATTAATGATGATGATGATGATGATGATGAT 249
QY 77 AspGlySerLeuCysIleLeuGluAlaLeuLysArgSer----- 89
250 GACGGTTCCTCTTGTATGAGAACTCTCAACAGCTTCTCACTGCAATATCATCAC 309
QY 90 GlnThrGlnValMetValProThrSerSerProLysLeuGluAspPheLeuGlyGlyAla 109
310 CATATGCAAGTT-----TCATCTCAAAAGATGAGAATTTT----- 345
QY 110 ThrMetGlyThrHisGluTyrGly---SerHisGluArgGly-----LeuSerLeuAsp 126
346 ---TTTGGACCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 402
QY 127 SerIleTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGln 146
403 AGTTATTTCTACAAATACATCATCATGCG-----CCAAACAAACAAACAACTTTCA 453
QY 147 ProPheArgGlnGlnGlyHisMetSerValGlnThrHisProTyrTyrSerGlyLeuAla 166
454 GAGTTC-----TTTAGCTTCCCTCAAACTGAAACCAACCAT----- 489
QY 167 CysHisGlyLeuTyrGlnAlaProLeuGlnGluGluThrThrLysGluThrHisValSer 186
490 -----GAGGAGAAACA-----AGAAACTTACAGAAAT 516
QY 187 AspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLysAsnTyrValAlaProThr 206
517 GAC-----CTTGTTTGACACATGGA----- 537
QY 207 ArgGluPheSerThrHisGlnGlnValLeuGluGlnGlnMetAsnCysGlyMetCysAsn 226
538 -----GAGGAGCTTTTAATGATGAGGATATATGAG 567
QY 227 GluArgAsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeu 246
568 GAATTTCA-----CAATGACTGAGCTTG 591
QY 247 SerMetSerProGlySerGlnSerSerCysValThrAla----- 259
592 TCCATGAGCCCTGGGTACAAATCTAGCTGATCACTGCTTCATCACCAACAAACAA 651
QY 260 -----ProSer 261
652 ACTCAAAACCAACAGCATCTCTGAAGCTTTGTGAGACAGATGCTGANTTTGAGACA 711

```





NAME: Baerian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-067200US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1680 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 190..1488  
 OTHER INFORMATION: /product= "APETALA 2 (AP2)"  
 US-08-700-152A-3

Alignment Scores:  
 Pred. No.: 1.58e-44 Length: 1680  
 Score: 517.00 Matches: 163  
 Percent Similarity: 43.76% Conservative: 65  
 Best Local Similarity: 31.29% Mismatches: 151  
 Query Match: 14.76% Indels: 142  
 DB: 2 Gaps: 19

US-10-024-632-2 (1-663) x US-08-700-152A-3 (1-1680)

QY 171 TyrglnAlaProLeuGluGluGluThrThryseGluThrHis 184  
 DB 214 CACCAACACAAAGAGAGAAAGATCTGAAGATTTGTTATTTCTTACCAAGTAAACGG 273  
 QY 185 -----ValSerAspCysSerSerLeuMetProGlnMetThrGluGluLeuLys--- 200  
 DB 274 GTTGATCTTTCTCTTAATTCAGCTCTTACAGTGTGTATCGAAGATGATCCGATGAC 333  
 QY 201 -----AsnThrValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu 217  
 DB 334 GATGAACCTTAACCGGCTGACACCAATTAACCACTTGTCAACCACTGATCTTCTCCGTGAG 393  
 QY 218 GlnGlnMetLeuAsnGlyMetGlyAsnGluArgAsnGlyValSer 232  
 DB 394 ATGATTTCTTAACGGCGGT-----GGTGTCTCTTCGCGCTTCCCTCGG 435  
 QY 223 -----LeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMetSer 249  
 DB 436 GCTCACTGGTTGGTGTAAATTTGT----- 462  
 QY 250 ProGlySerGlnSerSerCysValThrAlaProSerGlyThrAspSerValAlaValAsp 269  
 DB 463 -----CACTCGATCTAGCCACCGGATGCTCCGGGGTAAAGCTAACCAAGCTTCCGCC 513  
 QY 270 AlaLysLysArgGlyHisAlaLysLeuGlyGlnGlnProValHisArgLysSerIle 289  
 DB 514 GCT-----GCCGTAGTGAGCGCGACAGCCGCTTG---AAAAAGAGTCGG 555  
 QY 290 AsnThrPheGlyGlnArgThrSerGlnThrArgGlyValThrArgHisArgThrThyGly 309  
 DB 556 CGTGACCAAGATCAAGAGATTCAGATTAAGAGGTGATGATTTTACCGCGATCCGGA 615  
 QY 310 ArgTyrglnAlaHisLeuThrAspAsnSerCysLysLysGlnGluGlnThrArgLysGly 329  
 DB 616 AAGTGGGAATCTCAATTTGGGAC-----TGT-----GGG 645  
 QY 330 ArgGlnValTyrglnLeuGlyGlyTyrglnPheMetGluGluLysAlaAlaArgAlaTyrgln 349  
 DB 646 AAACAGATTACTAGTGGATGATTAACACTGCTACGACGCTCGAGCATTAATGATGA 705  
 QY 350 AlaAlaLeuLysTyrglnTyrglnProSerThrHisIleAsnPheSerIleGluAsnTyrgln 369  
 DB 706 GCTGCTAATTAATTCGCTGGAGTAGAAGCGGATATCAATTTCAACATGACGATTAATGAT 765

QY 370 ValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyrglnAlaHisLeuArgArg 389  
 DB 766 GATGACTTGAACAGATGACTTAATTTAACCAAGAGAGTTCGTACACGTACTTCGCCGA 825  
 QY 390 LysSerSerGlyPheSerArgGlyAlaSerIleTyrglnArgGlyValThrArgHisIleGln 409  
 DB 826 CAAGACACAGGCTTCCCTCGAGGAAGTTCGAATTAAGAGGTGATGATCTTGG---CATAG 882  
 QY 410 HisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrglnGly 429  
 DB 883 TGTGCTCTTGGGAAGCTCGATAGGCTCATATTTCTTAGCAAAAAGTATGATTTATTTGGCT 942  
 QY 430 ThrPheSerThrGlnGluGluAlaAlaGluAlaTyrglnPheValAlaAlaIleLysPheArg 449  
 DB 943 TTGTCGACACCGAGGTGGAAGCTGTGAGCTTACGATTAAGATGAAGCGCAATCAATGTAAC 1002  
 QY 450 GlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrglnPheValGluArgIleMetAla 469  
 DB 1003 GGCACAAAGCGCGTGAACCACTTGTATCCGATATTTCGATGAGGAATCAATGCCGAG 1062  
 QY 470 SerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLysLysAspAsnProArgAsn 489  
 DB 1063 TCATCAGGG-----AATCTTAATCTTCCACAAAT 1092  
 QY 490 LysAspIleAspTyrglnLysSerValAlaThrSerValAsnAsnGluGluThrValGln 509  
 DB 1093 CACAACTCGAT----- 1104  
 QY 510 ValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGluTrpLysMetValLeuPheAsn 529  
 DB 1105 CTGAGCTTGGAATATTCGGCTAATTCG-----AAG 1134  
 QY 530 HisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLysIleMetAsnGly 549  
 DB 1135 CATAAAGTCAAGATATGCGGCTCAGATGAACCAACAAACAGAT----- 1182  
 QY 550 GlyAsnTyrglnAsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAspSer 569  
 DB 1183 -----TCTCCCACTGAATGAAGTTCTTGATTAATGTCNA 1218  
 QY 570 ValGlySerGlyGlnHis-----AsnMetLeuAspLysSerIle 584  
 DB 1219 ACCGGAATGCTTAACCATTCCTCCAAATTCMAACCAACCAATTTCCGGGCGACGACAACT 1278  
 QY 585 GlyThr-----HisPheSer----- 589  
 DB 1279 GGTAGCGAGGCGGATTTCTCACTGTTCCGGCGCGTGAAGAACCCGGTTGATGGTCGG 1338  
 QY 590 -----AsnThrSerSerLeuValThrSerLeuSerSerSerArgGluAlaSerProGluLys 608  
 DB 1339 GCCTCGACCAACCAAGTGTGAACAAATGCTGAGATCTACAGATTTCTTCTCTCATCAT 1398  
 QY 609 ArgGlyProSerLeuLeuPheProMetProPheMetGluThrLysIleValAsnProIle 628  
 DB 1399 -----CAAAACAGATTTTAAATTTACT 1422  
 QY 629 GlyThrSerValThrSerTrpLeuProSerProThrValGln-----MetArgPro 645  
 DB 1423 TCTACTCTTCATCAAAATATGCTGCGACAAATGAGCTTCCAACTCTCTCTGAGACCT 1482  
 QY 646 Ser 646  
 DB 1483 TCT 1485

RESULT 6  
 US-09-026-039-3  
 : Sequence 3, Application US/09026039  
 : Patent No. 6328567  
 : GENERAL INFORMATION:  
 : APPLICANT: Jofuku, K. Diane  
 : APPLICANT: Okamoto, Jack K.  
 : TITLE OF INVENTION: Methods for Improving Seeds  
 : NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/026,039  
 FILING DATE: 19-FEB-1998  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/912,272  
 FILING DATE: 15-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/879,827  
 FILING DATE: 20-JUN-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/700,152  
 FILING DATE: 20-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Baetian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-067230US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 11721 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: -  
 LOCATION: 1..11721 /note="Arabidopsis APETALA2 (AP2)"  
 OTHER INFORMATION:  
 OTHER INFORMATION: genomic sequence"  
 US-09-026-039-3  
 Alignment Scores:  
 Pred. No.: 2,86e-19 Length: 11721  
 Score: 292.00 Matches: 139  
 Percent Similarity: 34.05% Conservative: 50  
 Best Local Similarity: 25.05% Mismatches: 121  
 Query Match: 8.34% Indels: 246  
 DB: 4 Gaps: 16  
 US-10-024-632-2 (1-663) x US-09-026-039-3 (1-11721)  
 QY 130 TTAAsenSerGlnAsnAlaGlnProAsnArgAspLeuLeuSerGlnProPheAr 149  
 Db 1004 TATTAGCTTCTTAACCTGAGAGATCATACAGAGATTTGAACCTTCAAG 1063  
 QY 149 ggGInGInGlyHisMetSerValGlnThrHisProCysTyrSerGlyLeuAlaCysHisG1 169  
 Db 1064 ATCAAAATCAAGAAACCAAAAACCAAAAATATGTGGAT 1106  
 QY 169 YLeuTYGlnAlaProLeuGlnGlnGlnThrThyGlySerGlnThrHis 184  
 Db 1107 -CTAAAGAGACGACCAACAAACAAAGAGAGAGATCTGAAGAGTTTGTATTTC 1165  
 QY 185 -----ValSerAspCysSerSerSerLeuMetProGlnMetTh 196  
 Db 1166 TTCACCAAGTAAACGGGTGGATCTTCTTAATCTTAAGCTTGTATTGGA 1225  
 QY 196 TGIUGLYLeuLys-----AsnTrpValAlaProThrArgGlnPheSerThrH1 212

Db 1226 AAGATGATCCGATGACGATGAACCTTAACCGGGTCAAGACCAATAACCACTTGCACCA 1285  
 QY 212 gGlnGlnValLeuGlnGlnGlnGlnMetLeuCysGlyMetC1YAsnGlnArgAsnGlyValSe 232  
 Db 1286 TCAAGTTCTTCCCTGAGATGATCTTAACGCGCGT-----GGTGTTC 1327  
 QY 232 r-----LeuGlySerValGlyCysGlyGlnLeuGlnSerLe 244  
 Db 1328 TTCTGAGCTTCTCTCGGCTCACTGGTTGGTGTAGTTTGT----- 1370  
 QY 244 uSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaProSerGlyThrAs 264  
 Db 1371 -----CACTCGGATCTAGCCACCGGATCTCGCGGTAA 1405  
 QY 264 pSerValAlaValAlaSerAlaValLeuValGlyHisAlaValLeuGlnGlnProVa 284  
 Db 1406 AGCTACCAACGTTGGCGCT-----GCCGTAGTACGCCGACACAGCCGTT 1450  
 QY 284 HisArgLysSerLeuAspThrPheGlyGlnArgThrSerGlnTrpArgGlyValThrAr 304  
 Db 1451 G---AAAAGAGTCCGCGCTGACAGTCAAGAGTCTAGATAGAGGTGTACGTT 1507  
 QY 304 gHisArgTrpThrArgLysArgLysGlnAlaHisLeuTrp-AspAsn----- 318  
 Db 1508 TTACGGCGGTACCGAGAGATGGAATCTCATATTGTGTAATATCTCATATTTTAATT 1567  
 QY 319 -----SerCysLysLeuGlnG 324  
 Db 1568 CGTAAATGATGTAATCTTAATTAATTAATTTTGTGTTGTTCTTCTGAAT 1627  
 QY 324 YGlnThrArgLysGlyArgGlnValYrLeuGlyGlyTr----- 337  
 Db 1628 TTCAGGAGACTG--CGGAACAAGTTTACTAGGAATTTATTTCTCATGTTT 1685  
 QY 337 ----- 337  
 Db 1686 TTGTATTGTTGGTGTGAATAATGTCATCAATAATTATTAATTAATCTGAATAGG 1745  
 QY 337 ----- 337  
 Db 1746 TGAGTTTGACACTGCTGATGACAGAGCTGGATTTTCTCTTGACTCTCTATAT 1805  
 QY 338 -----AspMetGlnGlu-LysAlaAla----- 344  
 Db 1806 TGAATGTTATTATTATTTTAAATAATACCGAAGAAATTTAATAAATTAATTTT 1865  
 QY 345 -----ArgAlaTrpAspLeuAlaAlaLeuLysTrpGlyProSer 358  
 Db 1866 AATTTGTTTATTAAATAGCATATGATAGAGCTGTATTAAATCCGAGAGTAGAA 1925  
 QY 359 ThrHisLeuAsnPheSerLeuGlnAsnTrpGlnValGlnLeuGlnGlu----- 374  
 Db 1926 GGGATATCAATTTCAACATCGAAGATTAATGATGATGATCTGAACAGGTAAATATAAT 1985  
 QY 374 ----- 374  
 Db 1986 TATAAATATATTGTTTATTAAAGATTTTAAAGGTTGGAGATTAATTTGAAT 2045  
 QY 375 -----MetLysAsnMetSerArgGlnGlnTrpValAlaHisLeuValArgLys 390  
 Db 2046 TGAATTTTATAGATACATTAATTAACCAAGAGAGAGTTCGACACGTAATCTCGCGCA 2105  
 QY 391 SerSerGlyPheSerArgGlyAlaSerLeuTrpArgGlyValThrArgHisGlnHis 410  
 Db 2106 AGCAGAGCTTCCCTCGAGAGAGTTCGAAGATAGAGGTGCACATTG---CATAGAGGT 2162  
 QY 411 G1YArgTrpGlnAlaArgLysLeuGlyArgValAlaGly----- 422  
 Db 2163 GGTCTTGGAGAGCTCGAATGGCTCAATTTTAAAGCAAAAGTAAATTTCTCTCATTTT 2222  
 QY 422 ----- 422





```

QY 338 -----AspMetGluGlu-LysAlaAla----- 344
Db 9916 TGAGTTGTTATTATTATTATTATTAAATACCGGAAAGAAATTTATATAATTAATTT 9857
QY 345 -----ArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyProSer 358
Db 9856 AATTTGTTTATTATTATAGAGCATATGATAGAGCTGCTATTAATTCGGAGAGTAA 9797
QY 359 ThrHisIleAsnPhseSerIleGluAsnTyrGlnValGlnLeuGluGlu----- 374
Db 9796 GCGGATATTCATTTCAACATTCGAAATATGATGATGACTTCAACAGCTAAATATAAT 9737
QY 374 ----- 374
Db 9736 TATTAACATAATTTGTTTTATTATTAACGATTTTAAAGTTGGAGATTAAATATGAAAT 9677
QY 375 -----MetLysAsnMetSerArgGlnGluTyrValAlaHisIleuArgLys 390
Db 9676 TGAATTTTATAGTACTAATTTTAAACAGAGAGAGTTCGTACAGTACTTCGCGACAA 9617
QY 391 SerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHis 410
Db 9616 AGCAGACAGCTCCCTCCGAGAGAGTTCGAAATATGAGGTGTCACATTG---CATAGTGT 9560
QY 411 GLYArgTrpGlnAlaArgIleGluArgValAlaGly----- 422
Db 9559 GGTGCTGGAGACTCGAATGGTCAATTCAGGCAAAAGTAAATTTCTCATTTT 9500
QY 422 ----- 422
Db 9499 ATATGACTCGAAACCTCATTTTAGTTGTATTAACTTGAGTTTGTCTCTG 9440
QY 423 -----AsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGluAlaAla-- 438
Db 9439 AATCTTATAAATAGCATGTATTATTGGTGTTCACACCGACGTCGAAGCTGCTAG 9380
QY 438 ----- 438
Db 9379 GTAAATGCTTTTGTGTTGATTCTACACACACATTTGTATTAATGTTTTCTCGTT 9320
QY 439 -----GluAlaTyrAspValAlaAlaIleLeu 447
Db 9319 ACTAATGATTTTCATTATTATTATATATATACAGAGTTTCGATAAGCTGCAATCA 9260
QY 447 sPheArgGlyAlaAsnAlaValaThrAsnPheserIleSerArgTyrAspValGluArgI 467
Db 9259 ATGTAACGGCAAGACCGCTGACCAACTTGTATCGAGTATTTCAGATGAGAA----- 9205
QY 467 eMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAla 479
Db 9204 -----CTCAATGCCGTAATTTGCT 9184

RESULT 8
US-09-313-294A-6727
; Sequence 6727, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laloudi, Raghnunach V.
; APPLICANT: Ico, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313, 294A
; NUMBER OF SEQ ID NOS: 1999-05-14
; SOFTWARE: PERL Program
; SEQ ID NO 6727
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature

```

```

; OTHER INFORMATION: Inocyte ID No. 6476212 700352302H1
; NAME/KEY: unsure
; LOCATION: 297
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6727

Alignment Scores:
Pred. No.: 6,4e-19 Length: 301
Score: 263.00 Matches: 54
Percent Similarity: 77.38% Conservative: 11
Best Local Similarity: 64.28% Mismatches: 19
Query Match: 7,518 Indels: 2
DB: 4 Gaps: 0

US-10-024-632-2 (1-663) x US-09-313-294A-6727 (1-301)
QY 332 ValTyrLeuGlyGlyTyrAspMetGluGluLysAlaAlaArgAlaTyrAspLeuAlaAla 351
Db 51 GCTTACTCGGGGCGCTATGACAGGAGAGGCAAGTCTGCTGCTATGACTCGAGCT 110
QY 352 LeuLysTyrTrpGlyProSerThrHisIleAsnPhseSerIleGluAsnTyrGlnValGln 371
Db 111 CCAAGTACTGGGCGCTCGAAGACTCTGCTCAACTTCCTCGGAGGATTACTCCAGGAG 170
QY 372 LeuGluGluMetLysAsnMetSerArgGlnGluTyrValAlaHisIleuArgLysSer 391
Db 171 ATGCCGAGATGAGAGCGCTGCTCCCGGAGAGTAACTGGCTCCCTCCCGCGAGAGC 230
QY 392 SerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGly 411
Db 231 AGCGGCTTCTCCGAGGCGGTCTCCAGTACAGAGGCGTCCG-AGGCATCACCAAGCGG 289
QY 412 ArgTrpGlnAla 415
Db 290 AG-TGGAGNCA 300

RESULT 9
US-09-198-119C-54
; Sequence 54, Application US/09198119C
; Patent No. 6417428
; GENERAL INFORMATION:
; APPLICANT: Thomasow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Ottosen, Kirsten
; APPLICANT: Gilmour, Sarah
; APPLICANT: Zhang, Daniel
; APPLICANT: Zhang, Cai-Zhong
; TITLE OF INVENTION: Plant Having Altered Environmental Stress Tolerance
; FILE REFERENCE: 19117, 713 Seq List
; CURRENT APPLICATION NUMBER: US/09/198, 119C
; NUMBER OF SEQ ID NOS: 1998-11-23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: bncbf5 gene

```

US-09-198-119C-54

Alignment Scores:

Pred. No.:	4,78e-07	Length:	898
Score:	158.50	Matches:	62
Percent Similarity:	37.00%	Conservative:	39
Best Local Similarity:	22.71%	Mismatches:	97
Query Match:	4.53%	Indels:	75
DB:	4	Gaps:	7

US-10-024-632-2 (1-663) x US-09-198-119C-54 (1-898)

```

QY 359 ThrHisIleAsnPhSerIleGluAsnTrpGlnValGlnLeuGluGluMetLysAsnMet 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 46 ACTTACTTACACTCAGCCTTTCACAGTTTCAAAAAAGATTTCACAGATGAATCAGTC 105
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 379 Ser-----ArgGlnGluTrpValAlaHisIleuArgArgLysSerSer 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 106 TCTACTTTTCTGAACTGCTCCGCTCGAGAACGAGCTCCGCTTATATACGAAAGTGAT 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 393 GlyPheSerArgGlyAlaSer----- 399
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 166 GATTACATTTTGGCGGCGAGCTGTCCACAGAACCTGCTGTAGAGAAAGTTTCAGAG 225
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 -----IleTrpArgGlyValThrArgHisIleGlnHisGlyArgTrpGlnAla 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 226 ACAGCCACCCCATTTTACAGAGAGTT---CGTCTGAGAAAGTCAGGTAAAGTGAGTGAT 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 416 ArgIleGlyArgValAlaGlyAsnLysAsnLeuTrpLeuGlyTrpPheSerThrGlnGlu 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 283 GAAGAGAGGAGAACCAACAGAAATCTAGAAATTGGCTCGGAACTTTCAAAACAGCTGAG 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 GluAlaIleGluAlaTrpAspValAlaAlaIleLysPheArgGlyAlaAsnAlaValThr 455
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 343 ATCGAGAGCTGTGCTCAGAGCTTGCCCTTAAAGTCCGTGAGAAAGGCGCTGCTC 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 456 AsnPhe----- 459
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 403 AACTTCGCGCACTCGGCTTGCGGCTCCGATCCCGAGACGACCTCGCCCAAGATATC 462
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 SerArg-----TyrAspValGluArgGlyLeuMetAlaSerSer 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 463 CAGAGAGGCTGCTCGAAGCCGCAATGGCTTTTGAAGCCGAGAAAGATGATACACAGC 522
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 472 AsnLeuLeuAlaGlyGluLeuAlaArgArgLysLysAspAsnAspProArgAsnLysAsp 491
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 503 AATGATCATGGCATGACATGAGCTTCTCAGGTGAGTTATGACACGACCGATCATGAC 582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 492 IleAspTrpAsnLysSerValValThrSerValAsnAsnGlnGlu----- 506
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 583 CTGGACATGAGAGACATGATGAGAGCTGTTTTCAGGAGGAAACAGAGAAAGGTTT 642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 507 -----ThrValGlnValGlnAlaGlyAsnAsnAsn 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 643 TACATGGCGAGAGAGACACGCTGTGTGGGTGTTGTTCCGAGAGAAACAGATAGCAAGGG 702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 AsnGluAsnAspSerGluTrpLysMet----- 525
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 703 TTTTACATGACAGAGAGATGATGTTCCGAGATCCGACCTTGTGGCTGATATGCGCGCA 762
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 526 ---ValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGln 544
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 763 GGGATGCTCTTACCCGCTGCTCCGTACATAGGAGACATATGATGATTCGAGAGAGAT 822
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 545 LysIleMetAsnCyGlyAsnTrpArgAsnSerAlaPhe 557
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 823 GCTGACATGAACTCTGGAATTATAGTACTCATATTTT 861
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10  
US-09-202-161B-6  
Sequence 6, Application US/09202161B  
Patent No. 6653533  
GENERAL INFORMATION:

APPLICANT: Purdue Research Foundation  
TITLE OF INVENTION: GENES ENHANCING DISEASE RESISTANCE IN PLANTS  
FILE REFERENCE: 7024-371  
CURRENT APPLICATION NUMBER: US/09/202,161B  
CURRENT FILING DATE: 1999-06-14  
PRIOR APPLICATION NUMBER: PCT/US97/10382  
PRIOR FILING DATE: 1997-06-12  
PRIOR APPLICATION NUMBER: 60/046,494  
PRIOR FILING DATE: 1997-05-14  
PRIOR APPLICATION NUMBER: 60/019,633  
PRIOR FILING DATE: 1996-06-12  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: ASCII  
SEQ ID NO 6  
LENGTH: 1405  
TYPE: DNA  
ORGANISM: Lycopersicon esculentum

US-09-161B-6

Alignment Scores:	1.67e-06	Length:	1405
Pred. No.:	156.50	Matches:	47
Score:	38.89%	Conservative:	16
Percent Similarity:	25.01%	Mismatches:	54
Best Local Similarity:	4.47%	Indels:	45
Query Match:	4	Gaps:	4

US-10-024-632-2 (1-663) x US-09-202-161B-6 (1-1405)

```

QY 359 ThrHisIleAsnPhSerIleGluAsnTrpGlnValGlnLeuGluGluMetLysAsnMet 378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 510 ACCGAGATCACTTATGCGCTGACCAATGATCGCGACAGAAACGAAAGATCGGAG 569
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 379 SerArgGlnGluTrpValAlaHisIleuArgArgLysSerSerGlyPheSerArgGlyAla 398
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 570 TCTCCGATTTCTGAGTCACT-----CGTCGAGAAAG----- 602
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 399 SerIleTrpArgGlyValThrArgHisIleGlnHisGlyArgTrpGlnAlaArgIleGly 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 603 -----TTTAAAGGGGT---CGTCAAGACCGTGGGCTGTTGGCTGCAAGATTCCG 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 ArgValAlaGlyAsnLysAsnLeuTrpLeuGlyTrpPheSerThrGlnGlnGlnAla 438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 654 GACCGACCCCGGAGAAACGGGTGTGGTGTGCTTATGACACCCAGAAAGACAGCT 713
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 439 GluAlaTrpAspValAlaAlaIleLysPheArgGlyAlaAsnAlaValThrAsnPheAsp 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 714 GTCGTTTACGATTAAGCTGCAAGTCTCAAGGCTCGACGCGCTTACCAATTTTCCG 773
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 459 IleSerArgTrpAspValGluArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeu 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 774 GATTCACAC----- 782
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 479 AlaArgArgLysLysAsnAspProArgAsnLysAspIleAspTrpAsnLysSerVal 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 783 -----ACGCGGAGATGA 794
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 499 ValThrSerValAsnAsnGlnGluTrpValGlnValGlnAlaGlyAsnAsnAsnGln 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 795 ACGGTACCGTTTACGAAACCGAAACCGAGTCTGTTGCCGACGCTGAGATATAAGCGAA 854
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 519 AsnAsp 520
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 855 AACGAT 860
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11  
US-09-026-039-1  
Sequence 1, Application US/09026039  
Patent No. 6329567  
GENERAL INFORMATION:  
APPLICANT: Jofuku, K. Diane  
APPLICANT: Okamoto, Jack K.  
TITLE OF INVENTION: Methods for Improving Seeds

NUMBER OF SEQUENCES: 103  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA: US/09/026,039  
 APPLICATION NUMBER: US/09/026,039  
 FILING DATE: 19-FEB-1998  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/912,272  
 FILING DATE: 15-AUG-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/879,827  
 FILING DATE: 20-JUN-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/700,152  
 FILING DATE: 20-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bastian, Kevin L.  
 REGISTRATION NUMBER: 34,774  
 REFERENCE/DOCKET NUMBER: 023070-067230US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1669 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: -  
 LOCATION: 1..1669 /note="canola APET1A2 (AP2) domain"  
 OTHER INFORMATION: containing (ADC) gene"  
 FEATURE3:  
 NAME/KEY: misc feature  
 LOCATION: 1..532  
 OTHER INFORMATION: /note="gene sequences 5' to start site"  
 US-09-026-039-1  
 Alignment Scores:  
 Pred. No.: 3.65e-06 Length: 1669  
 Score: 154.50 Matches: 86  
 Percent Similarity: 32.03% Conservative: 43  
 Best Local Similarity: 21.39% Mismatches: 95  
 Query Match: 4.41% Indels: 180  
 DB: 4 Gaps: 8  
 US-10-024-632-2 (1-663) x US-09-026-039-1 (1-1669)  
 QY 153 HAmSeSeValGlnThrHisProTyrSerGlyLeuAlaCysHISeglyLeuTyrGln 172  
 DB 551 CATATTTCTTGGCTTTGTTCTCTTTATTACAGAGAGATGTGAACCTTAACAC 610  
 QY 173 AlaPro-----LeuGluGluGluThrThyLysGluThrHisValSerAspCysSer 190  
 DB 611 TCACCTGATACACACCAAGATCCGACGCTAGATGAGAAACGGGCTGGAGATGTACCAATC 670  
 QY 191 LeuMetProGlnMetThrGluGlyLeuLysAsnTyrValAlaProThrArgGluPheSer 210  
 DB 671 TCTATGAGATCATGACACACAGCTGTCTGTCTCTCTCTCCCGGTACCCGGATATTATTTT 730

QY 211 ThrHisGlnIValLeuGluGlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGly 230  
 DB 731 TCCG--AATCAATCATGAAACAG----- 752  
 QY 231 ValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMetSerPro 250  
 DB 753 -----GAACTTCAGAAATATCTGGTCCGTATCACT 785  
 QY 251 GlySerGlnSerSerCysValThrAlaProSerGlyThrAspSerValAlaValAspAla 270  
 DB 786 AGAAACCACTCTTGTTCGGTCGATCTTCACGGGTCTGCT 827  
 QY 271 LysLysArgGlyHisAlaLysLeuGlyGlnLysGlnProValHisArgLysSerIleAsp 290  
 DB 828 ---CGTCGGAAACCTAGAGCTGAGATAACACACGCCGTA---AAAAGAGCCGACGT 881  
 QY 291 ThrPheGlyGlnArgThrSerGlnTyrThrArgGlyValThrArgHisArgTyrThrGlyArg 310  
 DB 882 GGTCTCTGCTACGGAGCTCTCAGTATAGAGAGTTACTTTTATTCAGAACCGAAGCA 941  
 QY 311 TyrGluAlaHisLeuTyrP----- 316  
 DB 942 TGGAGTCAATATTGGTAATTATTCTTAACCCAGATATACCAATACTATTA 1001  
 QY 316 ----- 316  
 DB 1002 TTACCTAATAGTAAATCATCAAAATACATGTTTCATTTCATTGAGCCAAATACCCTATT 1061  
 QY 317 -----AspAsnSerCysLysLysGluGlnThrArgLysGlyArgGlnValTyrLeu 334  
 DB 1062 GTTGTGTTTAAATATGTTTGAATCTTATGAGGAGCATGCG--GGAAACCAAGTGTACTTA 1120  
 QY 335 Gly----- 335  
 DB 1121 GGTATGATCATGTAAATGTTGTTCACAAACAGATCAATATCTATTGAAACTAAGTTGTG 1180  
 QY 335 ----- 335  
 DB 1181 TTGTGTCTGTCATTTTATATGATTTCTTCGACCAAAATAAGTTTATTATTCCTTA 1240  
 QY 335 ----- 335  
 DB 1241 TATTACTTTTGTACATATTCAGGTGATTTGACACACACATGCCGTGCTCGTATG 1300  
 QY 336 -----GlyTyrAspMetGluGluLys----- 342  
 DB 1301 TTTTACTCATCCAAATATGATCAATTAAGCAATCTAATTTCTTATTTTGTAAATTG 1360  
 QY 343 -----AlaAlaArg-AlaTyrAspLeuAlaI 351  
 DB 1361 CTGATATACAAATTAATTGGGTGGGTAACTGTTTGGGACACGCTGCTCAGATAGACCCG 1420  
 QY 351 AleuLysTyrTyrPglyProSerThrHisIleAsnPheSerIleGluAsnTyrGlnValG 371  
 DB 1421 AGTTAAGTTTAAAGGTATGATGCAGATTAATTAATTCATATTGAAGACTATGTGAGGA 1480  
 QY 371 nLeuGluGluMetLys----- 376  
 DB 1481 TTTGAACACGTTAAATATTATTATTGTAGATTCAACCAATTGACTTATGATTACT 1540  
 QY 377 -----AanMetSerArgGlnGluTyrVa 384  
 DB 1541 CGAACAATAAACAAATTAATTTGTTGACATGACGCAAGTTGACAAAGAAAGATTCT 1600  
 QY 384 lAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyVa 404  
 DB 1601 GCATGTCATTAGAGCAAAAGACAGCTGGTTTCCAAAGCAGCTTAAGATATAGAGTGT 1660  
 QY 404 lThr 405  
 DB 1661 CACT 1664  
 RESULT 12

```

US-09-198-119C-86
/ Sequence 86, Application US/09198119C
/ Patent No. 6417428
/ GENERAL INFORMATION:
/ APPLICANT: Thomashow, Michael
/ APPLICANT: Stockinger, Eric
/ APPLICANT: Jaglo-Olcosen, Kirsten
/ APPLICANT: Gilmour, Sarah
/ APPLICANT: Zarka, Daniel
/ APPLICANT: Jiang, Cai-Zhong
/ TITLE OF INVENTION: Plant Having Altered Environmental Stress Tolerance
/ FILE REFERENCE: 19117, 713 Seq List
/ CURRENT APPLICATION NUMBER: US/09/198,119C
/ PRIOR FILING DATE: 1998-11-23
/ PRIOR APPLICATION NUMBER: US 08/706,270
/ PRIOR FILING DATE: 1998-09-04
/ PRIOR APPLICATION NUMBER: US 09/018,233
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/017,816
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/018,235
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/017,575
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/018,227
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/018,234
/ PRIOR FILING DATE: 1998-02-03
/ NUMBER OF SEQ ID NOS: 95
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 86
/ LENGTH: 854
/ TYPE: DNA
/ ORGANISM: Brassica rapa
/ FEATURE:
/ OTHER INFORMATION: brcbf7 gene
US-09-198-119C-86

Alignment Scores:
Pred. No.: 5,02e-06 Length: 854
Score: 148.50 Matches: 53
Percent Similarity: 37.56% Conservative: 24
Best Local Similarity: 25.85% Mismatches: 75
Query Match: 4.24% Indels: 53
DB: Gaps: 5

US-10-024-632-2 (1-663) x US-09-198-119C-86 (1-854)
QY 400 ILETYRARGLYVALTHRAIGHISHISGLNHSGLYARGTTPGlnAlaArgIleGIYArg 419
DB 230 ATTTCAGAGAGATT---CGTCTTAGAAAGTCAGGTGGGTGTGTGAAGTGGAGAA 286
QY 420 VALAAGLYAsnlyAspleuTYrleuGIYThrPheSerThrGlnGluAlaArgIleGIYArg 439
DB 287 CCAAAAGAAATCTAGATTTGGCTCGAATCTTCAAAACAGCTGAGATCGCAGCTGCT 346
QY 440 AATATYAspValAlaAlaIleIysPheArgGlyAlaAsnAlaValThraPhe----- 457
DB 347 GCTCAGCAGAGCTGGCCGCTTACCTCCGTGGAAGAGCGCTGCTCACTGACCGAC 406
QY 458 -----AspIleSerArgTYAsp 463
DB 407 TCGGCTTGGCGGCTCCGATCCCGAGACACACCTGCCAAGATATCCAAAGAGCTGCT 466
QY 464 ValGluArgIleMetAla-----SerSerAsnleu 474
DB 467 GCTGAAGCCGCACTTGGCTTTGAGCGGAGAGAGTGTATCCACGACGACGAATGATCAT 526
QY 475 AAGAGlyleuAlaArgArgIysIysAspAsnAspProArgAsnIysAspIleAspTYr 494
DB 527 GCGATGACATGCTCTCTCAAGTTAGCTTAATGACAGACGAGATCAAGACCTGACATG 586
QY 495 AsnlySerValValThrSerValAsnAsnGlnGlu----- 506

```

```

DB 587 GAGGAGACAGTGTGGAGCGCTGTTTACGGAGACAGACAGAGCGTTTACATGACG 646
QY 507 -----ThrValGlnValGlnAlaGlyAsnAsnAsnGlnIleAsn 519
DB 647 GAGGAGACAGCGTGTGAGAGCTATTGTTCCGAGAGACAGATGACCAAGGTTTACATG 706
QY 520 AppSerGluTYrPheMet-----ValIleu 527
DB 707 GACGAGAGATGGATGTTGGAGATCCGACCTTGTGCTGATATGCGGACGAGATGCTC 766
QY 528 PheAsnHisProSerGlnGlnGlnAlaAsnIysAsnIysSerAspGlnIleMet 547
DB 767 TTACGCGCGCGCTCCGATCAATGGGAGACATATATGATGACTTCGAAGAGATGCTGACATG 826
QY 548 AencysGlyAsnTYr 552
DB 827 AACCTCTGGAATTAT 841

RESULT 13
US-09-198-119C-70
/ Sequence 70, Application US/09198119C
/ Patent No. 6417428
/ GENERAL INFORMATION:
/ APPLICANT: Thomashow, Michael
/ APPLICANT: Stockinger, Eric
/ APPLICANT: Jaglo-Olcosen, Kirsten
/ APPLICANT: Gilmour, Sarah
/ APPLICANT: Zarka, Daniel
/ APPLICANT: Jiang, Cai-Zhong
/ TITLE OF INVENTION: Plant Having Altered Environmental Stress Tolerance
/ FILE REFERENCE: 19117, 713 Seq List
/ CURRENT APPLICATION NUMBER: US/09/198,119C
/ PRIOR FILING DATE: 1998-11-23
/ PRIOR APPLICATION NUMBER: US 08/706,270
/ PRIOR FILING DATE: 1998-09-04
/ PRIOR APPLICATION NUMBER: US 09/018,233
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/017,816
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/018,235
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/017,575
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/018,227
/ PRIOR FILING DATE: 1998-02-03
/ PRIOR APPLICATION NUMBER: US 09/018,234
/ PRIOR FILING DATE: 1998-02-03
/ NUMBER OF SEQ ID NOS: 95
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 70
/ LENGTH: 950
/ TYPE: DNA
/ ORGANISM: Brassica oleracea
/ FEATURE:
/ OTHER INFORMATION: brcbf4 gene
US-09-198-119C-70

Alignment Scores:
Pred. No.: 6.8e-06 Length: 950
Score: 148.00 Matches: 53
Percent Similarity: 37.60% Conservative: 26
Best Local Similarity: 25.36% Mismatches: 78
Query Match: 4.23% Indels: 52
DB: Gaps: 5

US-10-024-632-2 (1-663) x US-09-198-119C-70 (1-950)
QY 400 ILETYRARGLYVALTHRAIGHISHISGLNHSGLYARGTTPGlnAlaArgIleGIYArg 419
DB 261 ATTTCAGAGAGATT---CGCCTTAGAAAGTGAAGTGGGTGTGTGAAGTGGAGAA 317
QY 420 VALAAGLYAsnlyAspleuTYrleuGIYThrPheSerThrGlnGlnAlaArgIleGIYArg 439

```



APPLICANT: Ratcliffe, Oliver  
 APPLICANT: Pilgrim, Marsha  
 APPLICANT: Jieng, Cai-Zhong  
 APPLICANT: Reuber, Lyne  
 TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES  
 FILE REFERENCE: MBI-010  
 CURRENT APPLICATION NUMBER: US/09/533,029  
 EARLIER APPLICATION NUMBER: 60/125,814  
 EARLIER FILING DATE: 1999-03-23  
 NUMBER OF SEQ ID NOS: 121  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 21  
 LENGTH: 1055  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 FEATURE:  
 OTHER INFORMATION: G19  
 US-09-533-029-21

Alignment Scores:  
 Pred. No.: 1.04e-05 Length: 1055  
 Score: 147.00 Matches: 67  
 Percent Similarity: 37.83% Conservative: 34  
 Best Local Similarity: 25.09% Mismatches: 100  
 Query Match: 4.20% Indels: 66  
 DB: 4 Gaps: 10

US-10-024-632-2 (1-663) x US-09-533-029-21 (1-1055)

QY 328 LysGlyArgGlnValTYrLeuGlyTYrAspMetGlnLysAlaAlaArgAlaTYr 347  
 DB 121 AAGGCGCGTAAATC-----ACGGCTGAGGAAGCTGCTGACAGCTCGAT 165  
 QY 348 AspleuAlaAlaLeuLysTYrTrpGly-----ProSerThr 359  
 DB 166 GCTTCGCGCGCGAGCACTTCTGGGTTTCTATTCCACTCCAACTCCATCCACCAAC 225  
 QY 360 HisLeuAspPheSerIleGluAsnTYrGlnValGlnLeuGlnLysMetSer 379  
 DB 226 CAAGTTAAAC-----GTGAAGAGAGGCGACGTGAAGAGAGCAG 264  
 QY 380 ArgGlnGluTYrValAlaHisLeuArgGlySerSerGlyPheSerArgGlyAlaSer 399  
 DB 265 GCAACAGAGCGGCGAAGCGAGAGAGAGAG-----AAT 300  
 QY 400 IleTYrArgGlyValThrArgHisIleGlnHisGlyArgTrpGlnAlaArgIleGlyArg 419  
 DB 301 GTTATTAGAGGGATA---CGTAAGCGTCATGGGAAATGGGCGCTGAGATTTCGAGAT 357  
 QY 420 ValAlaGlyAsnLysAspLeuTYrLeuGlyThrPheSerThrGlnGlnGluAlaGln 435  
 DB 358 CCACGAAAGGCTGTAAGTTGGCTTGTGTAACGCGGAGGAGGAGCTGCCATG 417  
 QY 440 AlaTYrAspValAlaAlaIleLysPheArgGlyValAlaAsnAlaValThrAspPheAspIle 459  
 DB 418 GCTTATTGATGTTGGCGGACCAAGATCCGTGTGATTAAGCCAGCTCACTTCCAGAT 477  
 QY 460 SerArgTYrAspValGluArgIleMetAlaSerSerAsnLeuAlaGlyGluLeuAla 479  
 DB 478 CTGCACCATCTCT-----CTTCCTCTAATTACTCTCCGCGCTCATCG 522  
 QY 480 ArgArgGlyLysAspAsnAspProArgAsnLysAspIleAspTYrAsnLysSerValVal 499  
 DB 523 CCACGATCAACCGATCAGCTCCGCGCAGAGAG-----GTC 558  
 QY 500 ThrSerValAsnAsnGlnGluThrValGlnValAlaGlyAsnAsnAsnGlnAsn 519  
 DB 559 TGGCGTTGTC-----TCTCAGAGT 576  
 QY 520 AspSerGluTYrPlyMetValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGly 539  
 DB 577 GAGAGCGAG-----TTAAGTCAACCGAGTTTCCCGGTGAGGTGTATAGGA 621

QY 540 AsnGlySerAspGlnLysIleMetAsnCyseGly-----AsnTYrArgAsn 554  
 DB 622 TTGGAAATGGGACGAGGTTTCAAACTGAGTTACGGAATTTAGCCGGAATTATGATCTG 681  
 QY 555 SerAlaPheSerMetAlaLeuGlnAspLeuIleGlyTleAspSerValGlySerGlyGln 574  
 DB 682 AAACAGCAGATATCGAGCTTGGAATCGTCTTGAGCTGACGGTAAACGCGCGAGCA 741  
 QY 575 HisAsnMetLeuAspGlnUser 581  
 DB 742 CCGAGTCACTTGAATGAGTCC 762

Search completed: March 13, 2004, 08:01:55  
 Job time : 131 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 13, 2004, 00:42:03 ; Search time 430 Seconds  
(without alignments)  
6550.129 Million cell updates/sec

Title: US-10-024-632-2  
Perfect score: 3502  
Sequence: 1 MKRINSENNITDDGNHNMWG.....RSPALISLHPVFAWTDT 663

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=x1h  
-O=/cgn2.1/USPTO/spec01/US10024632/runat 09032004 101207 10754/app query.fasta\_1.839  
-DB=N Geneseg 23Jan04 -QEXT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -STRAP=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdd  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10024632@cgn 1 1 352 @runat 09032004 101207 10754 -NCPU=6 -ICPU=3  
-NO MAP -LARGEJUTERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOCK  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
N\_Geneseg\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2002as:\*  
6: geneseqn2003as:\*  
7: geneseqn2003bs:\*  
8: geneseqn2003cs:\*  
9: geneseqn2003ds:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3502	100.0	2344	6	ABN84480
2	1715	49.0	2323	6	ABN84481
3	1408.5	40.2	1926	6	ABN84482
4	1349.5	38.5	2010	6	ABN84484
5	1338	37.9	1668	3	AAC43005
6	1328	37.9	2148	3	AA559220
7	1328	37.9	2148	7	ABX13437
8	1307	37.3	1738	3	AA559222

9	1104.5	31.5	1954	7	ACC00850	ACC00850 Glycine m
10	1059.5	30.3	2463	7	ACC00803	ACC00803 Glycine m
11	1048	29.9	2588	7	ACC00801	ACC00801 Glycine m
12	1032	29.2	1290	7	ACC00805	ACC00805 Glycine m
13	977	27.9	7367	6	ABN84483	ABN84483 Rice AINT
14	952.5	27.2	1874	7	ACC00799	ACC00799 Glycine m
15	951.5	27.2	1818	9	ADD30822	ADD30822 Plant yie
16	951.5	27.2	1818	9	ADD30822	ADD30822 Plant yie
17	951.5	27.2	1818	9	ADD30333	ADD30333 Plant yie
18	948	27.1	2011	4	ACC91399	ACC91399 Brassica
19	945.5	27.0	1500	9	ADD30559	ADD30559 Plant yie
20	944	27.0	2014	4	ACC91398	ACC91398 Brassica
21	938	26.8	1959	7	ACC00847	ACC00847 Catalpa s
22	936	26.7	1990	7	ACC00806	ACC00806 Glycine m
23	935.5	26.7	1935	7	ACC00848	ACC00848 Gossypium
24	930	26.6	2052	9	ADD30980	ADD30980 Plant yie
25	927	26.5	1824	7	ACC00802	ACC00802 Glycine m
26	922.5	26.3	1864	7	ACC00791	ACC00791 Zea mays
27	922.5	26.3	1878	7	ACC00846	ACC00846 Glycine m
28	920.5	26.3	2168	7	ACC00804	ACC00804 Glycine m
29	920	26.3	1758	7	ACC00794	ACC00794 Oryza sat
30	911.5	26.0	1753	7	ACC00792	ACC00792 Zea mays
31	894.5	25.5	2374	7	ACC00849	ACC00849 Plant yie
32	891.5	25.5	1941	9	ADD30349	ADD30349 Plant yie
33	718	20.5	551	7	ACC00796	ACC00796 Oryza sat
34	701	20.0	1726	7	ACC00845	ACC00845 Glycine m
35	696.5	19.9	1271	7	ACC00797	ACC00797 Glycine m
36	684	19.5	1753	7	ACC00844	ACC00844 Ricinus c
37	682	19.5	1465	6	ABQ81396	ABQ81396 Arabidops
38	682	19.5	1530	6	ABK65225	ABK65225 Arabidops
39	682	19.5	1530	6	AD31800	AD31800 DNA encod
40	682	19.5	1530	9	AD346784	AD346784 Thalecra
41	682	19.5	1530	9	AD330581	AD330581 Plant yie
42	681.5	19.5	1539	6	ABQ81395	ABQ81395 Arabidops
43	680	19.4	1619	7	ACC00843	ACC00843 Momordica
44	678	19.4	1710	7	ACC00785	ACC00785 Zea mays
45	674	19.2	1065	9	ADD31036	ADD31036 Plant yie

## ALIGNMENTS

RESULT 1	ABN84480	standard; cDNA; 2344 BP.
ID	ABN84480	
XX	AC	ABN84480;
XX	DT	21-OCT-2002 (first entry)
XX	DE	soybean AINTEGUMENTA-like polypeptide GmANT1 cDNA.
XX	KM	AINTEGUMENTA; ANT-like polypeptide; GmANT1; soybean; transgenic plant;
XX	OS	plant; Gene; se.
XX	XX	Glycine max.
XX	FT	Key
XX	FT	CD5
XX	FT	Location/Qualifiers
XX	FT	/*tag= a
XX	FT	/product= "GmANT1"
XX	XX	W0200259332-A2.
XX	XX	01-AUG-2002.
XX	PF	19-DEC-2001; 2001MO-US049294.
XX	PR	21-DEC-2000; 2000US-0257896P.
XX	PA	(MONS ) MONSANTO TECHNOLOGY LLC.
XX	PI	He SS, Dotson SB;
XX	XX	

DR MPI: 2002-599798/64.  
 DR P-PSDB; ABB79636.  
 XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in  
 PT improving agronomic, horticultural, and quality traits of plants, such as  
 PT increased size of plant organs.  
 XX  
 XX Example 1; Page 127-130; 169pp; English.

XX The present sequence is that of cDNA encoding GmANT1, a newly identified  
 CC AINTEGUMENTA-like (ANT-like) polypeptide of soybean. The cDNA comprises  
 CC the insert of clone CPR6763, which was obtained from a soybean sequence  
 CC database screening using Arabidopsis ANT as query. Also isolated was  
 CC clone CPR67626 (see ABBN4481) encoding GmANT2 (see ABB79637). GmANT1 and  
 CC GmANT2 show homology to ANT in 2 N-terminal Ap2 DNA binding domains, but  
 CC have C-terminal sequences that bear little, if any, homology to ANT. The  
 CC invention provides nucleic acids encoding ANT-like polypeptides  
 CC comprising, in the N-terminal to C-terminal direction, 2 Ap2 DNA binding  
 CC domains followed by an amino acid subsequence selected from those given  
 CC in ABB79632-35. ANT-like polypeptides have been identified in soybean,  
 CC rice, cotton and corn (see ABB79636-41). Nucleic acids (see ABBN4480-86)  
 CC encoding the ANT-like polypeptides can be used in the construction of  
 CC transgenic plants, especially corn, soybean, canola, wheat, cotton,  
 CC tomato or potato (all claimed) having improved agronomic, horticultural  
 CC or other quality traits, such as increased size of plant organs. These  
 CC plants are especially useful for production of ethanol or animal  
 CC feedstuff. The nucleic acids may also be used in antisense technology to  
 CC suppress endogenous ANT-like gene expression, in sense co-suppression to  
 CC modulate expression of endogenous ANT-like genes, as probes for  
 CC genetically and physically mapping the genes that they are part of, as  
 CC markers for traits linked to those genes, or in the identification of  
 CC loss of function mutant phenotypes of a plant due to a mutation in one or  
 CC more endogenous genes encoding the ANT-like polypeptides

XX Sequence 2344 BP; 738 A; 472 C; 548 G; 586 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No: 3,96e-284 Length: 2344  
 Score: 3502.00 Matches: 663  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-10-024-632-2 (1-663) X ABBN4480 (1-2344)

QY 1 MetLysArgLeuGlnGluSerAsnAspGluValAspAspGluValAsnAsnHisAsnTrpLeuGly 20  
 DB 242 ATGAAGGCGATTAATGAGTACACACCGATGATGAAACATCATACGTTGGGG 301  
 QY 21 PheSerLeuSerProHisMetLysMetGluAlaThrSerAlaAlaThrValProThrThr 40  
 DB 302 TTCTCTCTCCACCCCATGAAATGAGAGCTACTTACGACGCCCTGTTCCGACACC 361  
 QY 41 PheTyrMetSerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGly 60  
 DB 362 TTCTACATGTCCTCTTCACTCACTGTCCTCAACTTCGAAATGCTTACGGTCCGA 421  
 QY 61 GluAsnGlyAsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeu 80  
 DB 422 GAAATGGTAACCTTCATCTTCCACTTACGCTTATGCTCTCAAGCTCATATGGGTCATT 481  
 QY 81 CysLeuLeuGluAlaLeuLysArgSerGlnThrGluValMetValProThrSerSerPro 100  
 DB 482 TGAATCTTGAAGCTTCAAAAGATCCAAACCGCAAGTCAATGCTGCCAATTCGTCCTCG 541  
 QY 101 LysLeuGluAspPheLeuGlyAlaAlaThrMetGlyThrHisGluTyrGlySerHisGlu 120  
 DB 542 AAATTTGAGGAGCTTTCTAGGTTGGTGCACATATGGGAACTCAGCAATATGGAAGCCAGAG 601  
 QY 121 ArgGlyLeuSerLeuAspSerTyrTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsn 140  
 DB 602 AGAGGTTGAGCTTGAAGCATCTATTATTAATCCCAAAACGAGAGGCTCAACCCAAC 661

QY 141 ArgAspLeuLeuSerGlnProPheArgGlnGlnGlyHisMetSerValGlnThrHisPro 160  
 DB 662 AAGAGACCTTCTTTCACAAACCTTCAGGCAAGGCTCATAGAGTGTCCAAACACACCT 721  
 QY 161 TyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlnThrThr 180  
 DB 722 TATTACTAGGCTTGGCTTGGCTTGCATGTTATATCAAGACCGTTGGAGAAAGAAACA 781  
 QY 181 LysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLys 200  
 DB 782 AAGGAACCCACAGTGTGGATTCAGCTCCCTAATGCTCAATACAGAACGCTTGAAA 841  
 QY 201 AsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlnGlnMet 220  
 DB 842 AACTGGGTGCTCCCAACAGGAGATTTCATCCACGACGAGTTTGGAGCAGCAATG 901  
 QY 221 AsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGlyGlu 240  
 DB 902 AATTGTGGCATGGGAGATGAGAAATGCTGTCTTAAAGATCTGTGGGTGTGAGAG 961  
 QY 241 LeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaPro 260  
 DB 962 TTCAATCTCTAAGCTTATCTTATGACTCTGTTCTCACTAAGTTGTGTCTGCTCT 1021  
 QY 261 SerGlyThrAspSerValAlaValAspAlaLysLysArgGlyHisAlaLysLeuGlyGln 280  
 DB 1022 TCTGGAACAGATTCTTCTGCTGTGATGCAAAAGAGAGGCGCATGCTAACTTGCTCAG 1081  
 QY 281 LysGlnProValHisArgLysSerTyrLeuPheGlnGlyArgThrSerGlnTyrArg 300  
 DB 1082 AAGCAGCTGTGATGAAATATTCAGACATTTGGGCAAGAACCTCGCGATACAG 1141  
 QY 301 GlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTrpAsnSerCys 320  
 DB 1142 GGTGTCAACAGCATATGAGATGAGACTGTAGTATGAAGCCCATTTGGGATTAATGTTCC 1201  
 QY 321 LysLysGlnGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGln 340  
 DB 1202 AAGAGGAAGGGAACCTAGGAAGGACGACAGATATTTGGGGGTTATGATAGAG 1261  
 QY 341 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyProSerThrHis 360  
 DB 1262 GAGAAAGCTGCAAGAGCTTATGATCTCGGGCCCTTAAGTACGGGAGCACTTCAGCGCAT 1321  
 QY 361 IleAsnPheSerTyrLeuAsnTyrGlnValGlnLeuGlnGlyMetLysAsnMetSerArg 380  
 DB 1322 ATAAACCTTTCGATAGAAATTCACAACTTCACTTGAAGAAATGAAACATGAGCAGA 1381  
 QY 381 GlnGluTyrValAlaHisLeuArgLysSerSerGlyPheSerArgGlyAlaSerTyr 400  
 DB 1382 CAGGAATACCTTGCACCTTGAGAAAGAAAGACAGCGGGTTTCTAGAGGTGCTTCATATA 1441  
 QY 401 TyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgVal 420  
 DB 1442 TACAGAGGGGTTCACAAAGCATCACCAACAGAGATGGCAACGAGATAGGAGAGTT 1501  
 QY 421 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGlnGluAlaAlaGluAla 440  
 DB 1502 GCTGGGAACAAAGACCTTACCTTGGGAGCTTCAGCACCCAAAGGAGAGCAGCAAGCA 1561  
 QY 441 TyrAspValAlaAlaIleLysPheArgGlyValAsnAlaValThrAsnPheAspIleSer 460  
 DB 1562 TACATGTAGGGGAGTCAAAATTTGGGGGCAATGCAATCAACAACTTGAACATTCA 1621  
 QY 461 ArgTyrAspValGluArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArg 480  
 DB 1622 AGATACGATGTGGAGAAATCATGCGCAGTACGAAATCTCTCGCTGGGAGCTTGCAAG 1681  
 QY 481 ArgLysLysAspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValValThr 500  
 DB 1682 CGTAAGAAAGATTAACGATCTCTAGAAACAGAGCATAGACTTACAAACAGAGTGTAGTACA 1741



Accession	Gene	Protein	Length	Source	Notes
QY	SerValLeuAsnGluGluThrValGlnValGlnIleGlyAsnAsnAsnGluAsnAsp	520	ABN84481	standard; cDNA; 2323 BP.	
Db	1742 AGTGGAAACATGTAGGAAACGGTTCAGCTTCACGACGAAACACATTAAGAAACGAC	1801	ABN84481		
QY	SerGluTrpLysMetValLeuPheAsnHisProSerGlnGlnGlnIleAsnGlyAsn	540	ABN84481		
Db	1802 TCAGGTGGAAGATGGTTTATTATTTAACCAACCTTCACGACGACAAACGGCAATGGCAAT	1861	ABN84481		
QY	GlySerAspGlnIleValIleMetAsnGlyAsnTrpArgAsnSerPalaPheSerMetAla	560	ABN84481		
Db	1862 GGCACGTGACCAAAATATATGACTGTGAAATTTCAGAAACAGTGCATTTCTATGGCC	1921	ABN84481		
QY	LeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlu	580	ABN84481		
Db	1922 CTACAAAGTCTTATGGATGTATTGGTGGGTTCTGGGCGACGATATATGCTGACGAG	1981	ABN84481		
QY	SerSerIleValIleGlyThrHisPheSerAsnTrpSerSerLeuValThrSerLeuSerSer	600	ABN84481		
Db	1982 TCTACCAAAATTTGGAGCTCATTTTTCAAACAGTATGCGGTACACAGTTTAAAGCAGC	2041	ABN84481		
QY	SerArgGluAlaIleSerProGluLysArgGlyProSerLeuLeuPheProMetProMet	620	ABN84481		
Db	2042 TCAGAGAGGCTAGTCTCGAGAAAGGGGTCCTCGCTTCGTTCCATAGCTTCAGT	2101	ABN84481		
QY	GluThrIleValIleValAsnProIleGlyThrSerValThrSerTrpLeuProSerProThr	640	ABN84481		
Db	2102 GAAACAAAGATGTGTACACCCCATTTGTMACAGTGTACCTCTGGCTACCCCTCACCAAG	2161	ABN84481		
QY	ValGlnMetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrp	660	ABN84481		
Db	2162 GTTCAATAGAGCCTTCTCTGCTATCTTTGTCTCACTTCGAGTTTTCCTCTTGG	2221	ABN84481		
QY	661 ThrAspThr 663		ABN84481		
Db	2222 ACTGATACT 2230		ABN84481		
RESULT 2					
ABN84481	ABN84481	standard; cDNA; 2323 BP.			
XX	ABN84481;				
AC	21-OCT-2002 (first entry)				
DT					
DE	Soybean AINTEGUMENTA-like polypeptide GmANT2 cDNA.				
XX	AINTEGUMENTA; ANT-like polypeptide; GmANT2; soybean; transgenic plant;				
KM	plant; gene; ss.				
XX	Glycine max.				
OS					
XX					
PH	Key	Location/Qualifiers			
FT	CDS	25...2022			
FT		/*tag= a			
FT		/product= "GmANT2"			
XX					
XX	MO200259332-A2.				
PN					
XX					
PD	01-AUG-2002.				
XX					
PF	19-DEC-2001; 2001WO-US049294.				
XX					
XX	21-DEC-2000; 2000US-0257896P.				
PR					
XX	(MONS ) MONSANTO TECHNOLOGY LLC.				
PA					
PI	He SS; Dotson SB;				
XX					
DR	WPI; 2002-599798/64.				
XX	P-PSDB; ABB79637.				
PT	New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in				
XX	improving agronomic, horticultural, and quality traits of plants, such as				

PT increased size of plant organs.

PS Example 1; Page 134-137; 169pp; English

The present sequence is that of cDNA encoding GmANT2, a newly identified ANTIGENUNTA-like (ANT-like) polypeptide of soybean. The cDNA comprises the insert of clone CP87626, which was obtained from a soybean sequence database screening using Arabidopsis ANT as query. Also isolated was clone CP867663 (see ABN84480) encoding GmANT1 (see ABN76363). GmANT1 and GmANT2 show homology to ANT in 2 N-terminal AP2 DNA binding domains, but have C-terminal sequences that bear little, if any, homology to ANT. The invention provides nucleic acids encoding ANT-like polypeptides comprising, in the N-terminal to C-terminal direction, 2 AP2 DNA binding domains followed by an amino acid subsequence selected from those given in ABN79629-35. ANT-like polypeptides have been identified in soybean, rice, cotton and corn (see ABP76516-41). Nucleic acids (see ABN84480-86) encoding the ANT-like polypeptides can be used in the construction of transgenic plants, especially corn, soybean, canola, wheat, cotton, tomato or potato (all claimed) having improved agronomic, horticultural or other quality traits, such as increased size of plant organs. These plants are especially useful for production of ethanol or animal feedstuff. The nucleic acid may also be used in antisense technology to suppress endogenous ANT-like gene expression, in sense co-suppression to modulate expression of endogenous ANT-like genes, as probes for genetically and physically mapping the genes that they are part of, as markers for traits linked to those genes, or in the identification of loss of function mutant phenotypes of a plant due to a mutation in one or more endogenous genes encoding the ANT-like polypeptides

Sequence 2323 BP; 746 A; 482 C; 492 G; 603 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	5,51e-134
Score:	1715.00
Percent Similarity:	65.86
Best Local Similarity:	54.008
Query Match:	48.978
DB:	6
Length:	2322
Matches:	378
Conservative:	83
Mismatches:	165
Indels:	74
Gaps:	20

US-10-024-632-2 (1-663) X ABN84481 (1-2323)

QY 1 MetIyKrgTlEaNgIuSerNaSnrhThrApGjYlaNaSnHsAnTPlEugY 20  
 Db       :::  
 Db 25 ATGAAGAGTATGGAAAATGATGACAAATGCTGACTTAATATACAAAATGGTGGGT 84  
 QY       PheSerLeuSerProHisMet-----LysMetGluAlaThr 32  
 Db       85 TTTCACACTCTCTCTCAAATGCATATATAGSAGTTTTCACACTACAAACCTTCTCT 144  
 QY       33 SerAlaAlaThrValProThrThrPheTyrMetSerProSerGlnSerHisLeuSerAsn 52  
 Db       145 GCTCTCAAGTGGTCTCTCAACAGCTTTTCCACCACTGGTCCA-----CTTAGAGC 198  
 QY       53 PheGlyMetCysTyrGlyVal---GlyGluAsnGlyAsnPheHisSerProLeuThrVal 71  
 Db       159 TATGCTTTACATATGACTCTTAGCTGAAGTGAATGTTGGATTGTATACGCTTCCAAATC 258  
 QY       72 MetProLeuIuSerAspGlySerLeuCysIleGluGluAlaLeuLysArgSerGlnThr 91  
 Db       259 ATGCCCTCAAAATCTGATGGCTCTCTCTGTGATTGGAAACTTTATAGCAGGTCCAAAGCA 318  
 QY       92 GlnAlaMetValProHisSerSerProLysLeuGluAsnPheLeuGlyGlyAlaThrMet 111  
 Db       319 CAAGCAATGGCTACTTCAACACCAAAACTGGAAACTCTTAGTGGGGGAAGCATTG 378  
 QY       112 GlyThr---HisGlyIyr-----GlySerHisGluArgGlyLeuSerLeuAspSer 127  
 Db       379 GGAACCCCTCATCTACGAATGTAGTGGCCACAGAAACATGCTCTGAGCTTGGACAGT 438  
 QY       128 IleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnPro 147  
 Db       439 GTTTTATAC--ATCCAACTTCAGCGCGTGGACCAATATATATACCAAACTTACCAAAAC 495

QY 148 PheArgGln-----GlnGlyHisMetSerValGlnThrHisProTyrTyr 162  
 DB 496 CATGTTCAACACATTAGCACCAACCAACCAACAGAGGCTTCAACATATATAC 555  
 QY 163 SerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlnTyrThrVal 182  
 DB 556 TCTACCTTGAGAAACCATGATATATATTA-----GAAAGGCTCAAGCA 600  
 QY 183 ThrHisValSerAspCysSerSerLeuMetProGln----- 194  
 DB 601 AGCCAACTCTGACCAACCAATCTCATGTTCAAAACATGGGTGATGATGCCGT 660  
 QY 195 MetThrGlnGlyLeuLeuAsnTyrValAlaProThrArgGlnPheSer-----Thr 211  
 DB 661 CCTGTTCTGGCTTCAAGATTGG-----GAAGTGAAGAACTTCCAAAGCTAGCCATCA 714  
 QY 212 HisGlnGln-----ValLeuGlnGlnMetLeuPheCysGlyMetGlyAsnGlnArg 228  
 DB 715 CATGAGTCAAAGATGATTGTTCCATATGTGAGAGAAATGCTGGTGAATCAGAG- 768  
 QY 229 AsnGlyValSerLeuGlySerValGlyCysGlyGlnLeuGlnSerLeuSerMet 248  
 DB 769 -----TCCATTGATCATATGCTTATATGCTGACATGCTTGAAGCTTTCATG 819  
 QY 249 SerProGlySerGlnSerSerCysValThr-----AlaProSerGlyThr 263  
 DB 820 AGCTCTAGCTCTCACTCTAGACAGTGTCAAGTTCTCACGCTGCTCCTGCTGCTG 879  
 QY 264 AspSerValAlaValAlaSerAlaValAsnGlyValAsnGlyValGlnPro 283  
 DB 880 GATTCCTGTTGCCATGATACTAAGAAAGGGGCTGAAAGGTTGACCAACCAAT 935  
 QY 284 ValHisArgCysSerLeuAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThr 303  
 DB 940 GTTCAATGAGAGTCCATGATGACCTTTGGACAAAGAACTCCCATATAGAGATACA 999  
 QY 304 ArgHisArgTyrThrGlyArgTyrGlnAlaHisLeuTyrPheAsnSerCysLeuGln 323  
 DB 1000 AGGCATAGAGTGTGAGTGGAGATATGAACTCATCTTGGGACAAACAGCTGCAAGAAAGAG 1059  
 QY 324 GlyGlnThrArgGlyValArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnVal 343  
 DB 1060 GGACCAAGCAGAAAGAAAGAAAGCAAGTTTCTAGGGGGTTATGATGAAAGAAAGCT 1119  
 QY 344 AlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrProSerThrHisLeuAsn 363  
 DB 1120 GGAGAGCTTATGATCTAGCCGACCTCAAGTATTGGGGACCTCCACTCAATATACCTT 1179  
 QY 364 SerLeuGlnAsnTyrGlnValGlnLeuGlnMetLysAsnMetSerArgGlnGlnTyr 383  
 DB 1180 CCTTGGAAATATATATAAATGAACTTGAAGAAAGAAAGCAATGACTAGCAAGATAT 1239  
 QY 384 ValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGly 403  
 DB 1240 GTTGCTCATTTTGAAGAAAGAAAGCAAGATTTCTCAAGAGGGGCTTCCATGATACAGAGA 1299  
 QY 404 ValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsn 423  
 DB 1300 GTAAACAAGACACACACATGAGAGGTGCAAGCTCGAATGTGAGATGGCTGGAAAC 1359  
 QY 424 LysAspLeuTyrLeuGlyThrPheSerThrGlnGlnAlaGlnAlaGlyValTyrAspVal 443  
 DB 1360 AAAGATCTATATCTTGAACCTTTATACAAAGAGAGAGAGCTGAACCTATATAT 1419  
 QY 444 AlaAlaIleLysPheArgGlyValAlaAsnAlaValThrAsnPheAspIleSerArgTyrAsp 463  
 DB 1420 GCTGCTATATAATTCAGAGAGCGAATGCTGTACCAACTTGAACCTCAACAAATATGAT 1479  
 QY 464 ValGlnArgIleMetAlaSerSerAsnLeuAlaGlyGlnLeuAlaArgArgLys 483  
 DB 1480 GTGAGAAATATCATGACCAACCAACCTCTTACAGAGTGAAGTACTGAGGCAACGCA 1539  
 QY 484 AspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValValThrSerVal--- 502

DB 1540 GAGACGACCAATGAACACTCAGTCATTCATCAAAATCAATCAATTAAGCTTCTGATATGAG 1599  
 QY 503 AsnAsnGlnGlnThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnAspSerLeu 522  
 DB 1600 GACACTCAAGAGCACTATCTTAATGCAACGAGAGCTGTAGAGCGAAATGAT---CAG 1656  
 QY 523 TrpLysMetValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySer 542  
 DB 1657 TCGAAGATGCTTCTTAC---CAATCTCTAGCAACATGAGCAATCCACCAACAT 1713  
 QY 543 AspGlnLysIleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGln 562  
 DB 1714 GAG-----AGTACAGAACTAACCAAGTCTCTGCAAGTGGCTTGGAGC 1755  
 QY 563 AspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSer 582  
 DB 1756 AACATGTT-----CATCAAGAGATGAGAGATCAAGT 1788  
 QY 583 LysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArg 602  
 DB 1789 AAGCGAGAGACCATGTGTCAATCTCTTCAATGGCCACAAAGTTTGAAGCATCAAGA 1848  
 QY 603 GluAlaSerProGlnLysArgGlyProSerLeuLeuPheProMetProProMetGlnThr 622  
 DB 1849 GAAGGTAGCCCTGATATGACAAAGCTTGGCAATGCTCTCTGGAATGCTTCAACATGATCA 1908  
 QY 623 LysIleValAsnProIleGlyThrSerValThrSerTyrPheProSerProThrValGln 642  
 DB 1909 AAACATATGGCTACTATATCAATTAACGTAATCTTGGAGCCCTTCAACCCATTTGAGG 1968  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerThrPheAsp 662  
 DB 1969 -----CCAGCACTTACTTGTGCTCTCAAAATGCCAGTTTTCAGACTTGGACAGAT 2016  
 RESULT 3  
 ID ABN84482 standard; cDNA; 1926 BP.  
 XX AC ABN84482;  
 XX 21-OCT-2002 (first entry)  
 XX DT  
 XX DE  
 XX Rice AINTEGUMENTA-like polypeptide OsANTI cDNA.  
 XX KW AINTEGUMENTA; ANT-like polypeptide; OsANTI; rice; transgenic plant;  
 XX KM plant; gene; ss.  
 XX OS  
 XX Oryza sativa.  
 XX OS  
 XX Key  
 XX FT CDS  
 XX Location/Qualifiers  
 XX 1..1926  
 XX /\*tag= a  
 XX /partial  
 XX /product= "OsANTI"  
 XX /note= "the CDS does not include a stop codon"  
 XX  
 XX WC020259332-A2.  
 XX  
 XX 01-AUG-2002.  
 XX  
 XX 19-DEC-2001; 2001MO-US049294.  
 XX PF  
 XX 21-DEC-2000; 2000US-0257896P.  
 XX PR  
 XX (MONS ) MONSANTO TECHNOLOGY LLC.  
 XX PA  
 XX He SS, Dotson SB;  
 XX FI  
 XX WPI; 2002-595798/64.  
 XX DR P-PSDB; ABN79638.  
 XX XX  
 XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in

PT improving agronomic, horticultural, and quality traits of plants, such as  
 PT increased size of plant organs.

Example 2; Page 140-143; 169pp; English.

CC The present sequence is the coding region for OsANT1, a newly identified  
 CC ANTEGMENTA-like (ANT-like) polypeptide of rice. The cDNA was obtained  
 CC by PCR amplification using gene-specific primers (see ABN84496-97) using  
 CC particle cDNA as template. The cDNA was not amplified from root or leaf  
 CC cDNAs, suggesting tissue-specific expression. Initial OsANT1 clones had  
 CC been identified in a rice database screening using the newly identified  
 CC soybean ANT-like polypeptide GmANT1 as query. An additional rice clone  
 CC (see ABN84488) encoding OsANT2 (see ABN79639) was also isolated. OsANT1  
 CC polypeptide shares high homology with Arabidopsis ANT, GmANT1 and GmANT2  
 CC at the N-terminal AP2 DNA binding domains, shares conserved segments at  
 CC the N-terminus, and shares conserved segments with GmANT1 and GmANT2, but  
 CC not with ANT, at the C-terminus. The invention provides nucleic acids  
 CC encoding ANT-like polypeptides comprising, in the N-terminal to C-  
 CC terminal direction, 2 AP2 DNA binding domains followed by an amino acid  
 CC subsequence selected from those given in ABN79629-35. ANT-like  
 CC polypeptides were identified in soybean, rice, cotton and corn (see  
 CC ABN79635-41). Nucleic acids (see ABN84480-86) encoding the ANT-like  
 CC polypeptides can be used in the construction of transgenic plants, (all  
 CC especially corn, soybean, canola, wheat, cotton, tomato or potato (all  
 CC claimed) having improved agronomic, horticultural or other quality  
 CC traits, such as increased size of plant organs. These plants are  
 CC especially useful for production of ethanol or animal feedstuff. The  
 CC nucleic acids may also be used in antisense technology to suppress  
 CC endogenous ANT-like gene expression, in sense co-suppression to modulate  
 CC expression of endogenous ANT-like genes, as probes for genetically and  
 CC physically mapping the genes that they are part of, as markers for traits  
 CC linked to those genes, or in the identification of loss of function  
 CC mutant phenotypes of a plant due to a mutation in one or more endogenous  
 CC genes encoding the ANT-like polypeptides

XX Sequence 1926 BP; 438 A; 561 C; 585 G; 342 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,39e-108	Length:	1926
Score:	1408.50	Matches:	340
Percent Similarity:	55.05%	Conservative:	63
Best Local Similarity:	46.45%	Mismatches:	152
Query Match:	40.22%	Indels:	177
		Gaps:	26

US-10-024-632-2 (1-663) x ABN84482 (1-1926)

QY 13 GlyAsnAsnHisAsnTrpLeuGlyPheSerLeuSerProHisMetGluAlaThr 32  
 DB 13 GGGCGACAGCAACTGGTTAGGCTTCGCTCCCGACATGCGCGCATGAGAGTG 72  
 QY 33 SerAlaAlaThrValProThr----- 40  
 DB 73 CGGTCCTCTGAGCCATCGACTGCTCATCATCATCATCATCATCATCATCAT 132  
 QY 41 -----Phe 41  
 DB 133 GCTGCTGCTGCTGCCGAGCCATGCTCTCTCCGACAGCGGACAGACTGCAATTC 192  
 QY 42 TyrMetSerProSerGlnSerHisLeu-----SerAsnGly 54  
 DB 193 CTCTTCTCCCTCTGCGACAGATGCTGCTCTCTCACTGCTACTACTAGTGGC 252  
 QY 55 MetCysTrpGlyValGlyGluAsn-----GlyAsnPheHisSerProLeuThrValMet 72  
 DB 233 GCGCGCTACGAGACGAGACGACACCGCGCGCTCTACTACTGACACTCCCTGATG 312  
 QY 73 ProLeuYsSerAspGlySerLeuGlySerLeuGluAlaLeuYsAspSerGlnThrGln 92  
 DB 313 CCAACCAAGTCGATGCTCTCTCTGCAATCATGAGAGC----- 351  
 QY 93 ValMetValProThrSerProLeuGluAspPheLeuGlyValAlaThrMetGly 112

DB 352 ---ATGATGCGCTGCTGATCCGACAAAGCTCCAGAGACTTCTGGGGTGTGGCAATGAGCAT 408  
 QY 113 ThrHisGlyTrpGlySerHisGluArgGlyLeuSerLeuAspSerIleTyrTrpAsnSer 132  
 DB 409 GGGCAT-----GACCCCGGACCACTACTATATGCG 435  
 QY 133 GlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPheArgGlnGln 152  
 DB 436 CAGGGCCAAAGAACAGAGAT-----GGAAGAGGGCGGCGCTACCAAGAC 480  
 QY 153 HisMetSerValGlnThrHisProTyrTrpSerGlyLeuAlaCysHisGlyLeuTyrGln 172  
 DB 481 CACCGAGTACTG-----CCCTAACAC-----TACCAG 507  
 QY 173 AlaProLeuGluGluGluThrThrHisGluThrHisValSerAspCysSerSerLeuMet 192  
 DB 508 CCATTGACGAGAGCAGAGATGCTCAAGAGCGCCGACGCGCCCAATGAGAGCCCATG 567  
 QY 193 ProGlnMetThrGluGlyLeuYsAsnTrpValAlaProThrArgGluPheSerThrHis 212  
 DB 568 -----GCGGCGGCGCAAGAACTTC----- 585  
 QY 213 GlnGlnValLeuGluGlnGlnMetAsnGlyMetGlyAsnGluArgAsnGlyValSer 232  
 DB 586 -----CTGTCACACAGCTACGCGCGCTCTACGGACACAGAGATGCGC----- 630  
 QY 233 LeuGlySerValGlyCysGlyGlyLeuGlnSerLeuSerLeuSerMetSerProGlySer 252  
 DB 631 -----CAGCGGCTCAGCCCTCTCATGACCCCAAGGTTC 663  
 QY 253 Gln---SerSerCysValThr---AlaProSerGlyThrAspSerValAlaValAspAla 270  
 DB 664 CAGTCACAGAGCTGCTGATGTCACACTCCCAAGCAGCATCAGACAGATGCGGTGCTGCT 723  
 QY 270 ----- 270  
 DB 724 GAGCTGCTGCTGCTGATGCGGAGGAAAGCAACATTAATGACGTTGGGAGCAGCGT 783  
 QY 271 -----LysValArgGlyHisAlaLeuLeuGlyGlnGlnProValHisArgLysSer 288  
 DB 784 GTGCGAAGAAAGAGGCGACCGGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 843  
 QY 289 IleAspThrPheGlyGlyArgThrSerGlnTyrArgGlyValThrArgHisArgTrpThr 308  
 DB 844 ATTGACACGTTTGGGACAGAGCATCGAGTTCAGTTCAGGCGGCGGCGGCGGCGGCGG 903  
 QY 309 GlyArgTrpGluAlaHisLeuTrpAspAsnSerCysLysLysGlyGlnThrArgLys 328  
 DB 904 GGAAGATATGAAGCCCACTCTGGGATTAAGTTCAAAAAGATGAGACAGACAGAGAG 963  
 QY 329 GlyArgGlnValTyrLeuGlyGlyTyrAspMetGluGlnValAlaArgAlaTyrAsp 348  
 DB 964 GGAAGGCAAGTATCTAGTGGTGTATACCTGAAAGTAAAGCTGCGAGGGCTTATGAT 1023  
 QY 349 LeuAlaAlaLeuLysTyrTrpGlyProSerThrHisIleAsnPheSerIleGluAsnTyr 368  
 DB 1024 CTGGCTGGCTGAATATCTGGGGGCTATCTACGCAATATAATTTCCGTTGAAATACTAC 1083  
 QY 369 GlnValGlnLeuGluGlnMetLysAsnMetSerArgGlnGlnTyrValAlaHisLeuArg 388  
 DB 1084 CGAAGTGAATGAAGAGATGAAAGATGACAAAGGCAAGATATGTTGGACCTTGAGA 1143  
 QY 389 ArgLysSerSerGlyPheSerArgLysIleSerIleTyrArgGlyValThrArgHisHis 408  
 DB 1144 AGGAGAGAGAGGGGTTCTCTGCGGTCTTCACTACCGGGAGGTAAACAAGCATCAC 1203  
 QY 409 GlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 428  
 DB 1204 CAGCATGAGAGATGCAAGCTCGATGCGAGGGTGTGCGCAACAAGACCTGTATCTC 1263  
 QY 429 GlyThrPheSerThrGlnGluGluAlaAlaGluAlaTyrAspValAlaAlaIleLysPhe 448  
 DB 1264 GGCACCTTCAGACCTCAAGAGAGACACAGAGGATACGACATTCGCTGCCATCAAGTTC 1323



QY 87 s-----ArgSerGlnThrGlnValMetValProThrsSerProLysLeuGluAspH 105  
 Db 345 GGGAGCGCAACCAAGCAAGAGGAGTGTCTGGCGCTCCGCCCACTGAGAGATT 404  
 QY 105 eLeuGly---GlyAlaThrMetGlyThrHisGlyTyrGlySerHisGlyValGlyLeuSe 124  
 Db 405 CCAAGCGCGGCGCCCGGCGATG-----GCCCTGAG 434  
 QY 124 rLeuAsp-----SerLeuTyrTyrAsnSerGlnAsnAlaGlnProAsnArgAs 142  
 Db 435 CCAAGCAACCTCGGCTTCTACTAGCGCGCGCCACCGTCAACC----- 477  
 QY 142 rLeuLeuSerGlnProPheArgGlnGlnGlnMet-SerValGlnThrHisProTyrT 162  
 Db 478 -----CAAGGACACCGCCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527  
 QY 162 rSerGlyLeuAlaCysHisGlyLeuTyrGln---AlaProLeuGlnGlnGlnGlnGln 181  
 Db 528 CGGCGCGCGGCTTCTGCAAGTGGCGTGCATCCCGCGCGCGCGCGCGCGCGCGCG 587  
 QY 181 yseGlnThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlnGlyLeuYsa 201  
 Db 588 GCTG-GTGACGACACGATCCCGCGCGAGTg-----GCCGCCG 625  
 QY 201 snrTyrAlaProThrArgGlnPheSerThrHisGlnGlnValLeuGlnGlnMetA 221  
 Db 626 GCTGGCGCGCG-----ATGCAC----- 642  
 QY 221 snrCysGlyMetGlyLeuGlnArgGlnGlyValSerLeuGlySerValGlyCys----- 238  
 Db 643 -----GGCGCGCGCTTACGACATCGCCAAAGCGCGCGCGCGCGCGCGCGCG 694  
 QY 239 -----GlyLeuGlnSerLeuSerLeuSerMet---SerProG 251  
 Db 695 GCCCATCATCCCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 754  
 QY 251 yseGlnSerSerCysValThr-----AlaProSerGlyThrAspSerValAlaV 268  
 Db 755 GGTCCAGATCCAGCTGCGTCACTGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 814  
 QY 268 alaAspAla-----LysLysArgGlyHisAla---LysLeuGlyGlnLysGlnProValH 285  
 Db 815 TCGACCGCGCGAGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 874  
 QY 285 lAsArgLysSerTlleAspThrPheGlyGlnAlaGlnThrSerGlnTyrArgGlyValThrArgH 305  
 Db 875 ACCGCAAGTCCATGACAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 934  
 QY 305 lAsArgThrPheGlyArgTyrGlnAlaHisLysLysProAsnSerCysLysLysGlnGlyG 325  
 Db 935 ATAGTGGAGCTGGAGATATGAGGACACCTCTGGGACAAACACTCGCAAGAGAGAGG 994  
 QY 325 lnrThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaAlaA 345  
 Db 995 AGACCGAAG 1054  
 QY 345 rgaLAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHisLleAsnPheSerI 365  
 Db 1055 GGGGATATGATCTTGTCTGCGTCAAGTACTGGGCGCTTCACGCAATCAATCTCCGT 1114  
 QY 365 lLeuGlnTyrGlnValGlnLeuGlnGlnGlnMetLysAsnMetSerArgGlnGlnTyrValA 385  
 Db 1115 TCGAGGAGCTCCAG 1174  
 QY 385 lAsHisLeuArgArgLysSerSerGlyPheSerArgGlyValAserTlleTyrArgGlyValT 405  
 Db 1175 CTCACCTCAG 1234  
 QY 405 hrArgHisHisGlnHisGlyTyrGlnAlaAlaGlnLleGlyArgValAlaGlyValSerLysA 425  
 Db 1294

Db 1295 ACCTTTCTTGGGAGACATTCATCCGCTGCGCTTTTCCCGCGCGCGCGCGCGCGCATG 1354  
 QY 432 --SerThrGlnGlnGlnAlaAlaGlnAlaTyrAspValAlaAlaLleLysPheArgGlyA 451  
 Db 1355 CCGGCAAGCAG 1414  
 QY 451 lAsnAlaValAlaThrAsnPheAspLleSerArgTyrAspValGlnArgLlleMetAlaSerS 471  
 Db 1415 TCAACGCGCTCACCACTTCGATCAGAGAGTACAGCTGAGCAAGATCTCTGAGAGCA 1474  
 QY 471 eAsnLeuLeuAlaGlyGlnLeuAlaArgArgLys---LysAspAsnAspProArgAsnL 490  
 Db 1475 GCAGCTCTCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1533  
 QY 490 ysaAspLleAspTyrAsnLysSerValAlaThrSerValAsnAsnGlnGlnThrValGlnV 510  
 Db 1534 -----GCCGCGCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1567  
 QY 510 aGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGlnTyrLysMetValLeu----- 527  
 Db 1568 TCGAGCGCGCG-----AACGTGGCGAGTGGAGAGTGGCCACCGCGCGCGCG 1612  
 QY 528 -----PheAsnHisLeuProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGln 545  
 Db 1613 CGCTGCCAGCGCGCGCGCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1671  
 QY 545 yseLleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuI 565  
 Db 1672 -----CACCATGACCTCTCGCGCGAGCGAGCGCTTCTCGGTG---CTGAGAGACATCG 1720  
 QY 565 lGlyLys---LleAspSerValGlySerGlyGlnHisAsnMetLeuAspLysSerSerLysI 584  
 Db 1721 TGTGACCGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1755  
 QY 584 lGlyLysThrHisPheSerAsnThrSerSerLeuValAlaThrSerLeuSerSerArgGlnA 604  
 Db 1756 --GGCGCGCAATGCTCATGCGCGG-----ACAGGCTGGGCAACTCTCGGAGAC 1804  
 QY 604 lAserProGlnLysArgGlyPro-----SerLeuLeuP 615  
 Db 1805 AGAGCGCTGAC---AGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1861  
 QY 615 heProMetProProMetGlnThrLysLleValAsnProLleGlyThrSerValThrSerT 635  
 Db 1862 TCGCCAGCGCGCGCGCGCGCGCGCGCGAGACTGTACAGCCCGGTG-----CCGCTGAAACCT 1915  
 QY 635 rPLeu---ProSerProThrValGlnMetArgProSerProAla---LleSerLeuSerH 653  
 Db 1916 GGGCGCTGCGCTTCCCGCGCGGTAGCTCGGTGCGAGGAGCGCGCGCGCGCGCG 1975  
 QY 653 lAsnProValPheAlaSerTyrThrAsp 662  
 Db 1976 ACCTGCCAATGTTGCGCGCTGAGACCGAC 2004  
 RESULT 5  
 AAC43005  
 ID AAC43005 standard; DNA; 1668 BP.  
 AC AAC43005;  
 DT 17-OCT-2000 (first entry)  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37674.  
 XX Hybridation assay; genetic mapping; gene expression control;  
 XX protein identification; signal transduction pathway; metabolic pathway;  
 KM promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 OS  
 PN EPI033405-A2.

Wed Mar 17 08:18:55 2004

us-10-024-632-2.rng

Page 8

[illegible]

PR	06-UTL-1999	99US-0142390DP
PR	08-UTL-1999	99US-0146803DP
PR	09-UTL-1999	99US-0145920DP
PR	12-UTL-1999	99US-0143927DP
PR	13-UTL-1999	99US-0143542DP
PR	14-UTL-1999	99US-0143654DP
PR	15-UTL-1999	99US-0144005DP
PR	16-UTL-1999	99US-0144085DP
PR	19-UTL-1999	99US-0144086DP
PR	19-UTL-1999	99US-0144332DP
PR	19-UTL-1999	99US-0144331DP
PR	19-UTL-1999	99US-0144333DP
PR	19-UTL-1999	99US-0144334DP
PR	20-UTL-1999	99US-0144335DP
PR	20-UTL-1999	99US-0144332DP
PR	20-UTL-1999	99US-0144884DP
PR	21-UTL-1999	99US-0144881DP
PR	21-UTL-1999	99US-0145086DP
PR	22-UTL-1999	99US-0145085DP
PR	22-UTL-1999	99US-0145087DP
PR	22-UTL-1999	99US-0145089DP
PR	22-UTL-1999	99US-0145192DP
PR	23-UTL-1999	99US-0145145DP
PR	23-UTL-1999	99US-0145218DP
PR	23-UTL-1999	99US-0145624DP
PR	26-UTL-1999	99US-0145276DP
PR	27-UTL-1999	99US-0145213DP
PR	27-UTL-1999	99US-0145182DP
PR	27-UTL-1999	99US-0145919DP
PR	28-UTL-1999	99US-0145951DP
PR	02-AUG-1999	99US-0146386DP
PR	02-AUG-1999	99US-0146388DP
PR	03-AUG-1999	99US-0146389DP
PR	04-AUG-1999	99US-0147038DP
PR	04-AUG-1999	99US-0147204DP
PR	05-AUG-1999	99US-0147302DP
PR	05-AUG-1999	99US-0147192DP
PR	06-AUG-1999	99US-0147260DP
PR	06-AUG-1999	99US-0147303DP
PR	09-AUG-1999	99US-0147416DP
PR	09-AUG-1999	99US-0147393DP
PR	09-AUG-1999	99US-0147935DP
PR	11-AUG-1999	99US-0148171DP
PR	11-AUG-1999	99US-0148319DP
PR	12-AUG-1999	99US-0148341DP
PR	13-AUG-1999	99US-0148563DP
PR	13-AUG-1999	99US-0148684DP
PR	16-AUG-1999	99US-0149176DP
PR	17-AUG-1999	99US-0149175DP
PR	18-AUG-1999	99US-0149422DP
PR	20-AUG-1999	99US-0149426DP
PR	20-AUG-1999	99US-0149723DP
PR	20-AUG-1999	99US-0149929DP
PR	23-AUG-1999	99US-0149932DP
PR	23-AUG-1999	99US-0149930DP
PR	25-AUG-1999	99US-0150566DP
PR	26-AUG-1999	99US-0150684DP
PR	27-AUG-1999	99US-0151065DP
PR	27-AUG-1999	99US-0151065DP
PR	27-AUG-1999	99US-0151080DP
PR	27-AUG-1999	99US-0151103DP
PR	31-AUG-1999	99US-0151103DP
PR	31-AUG-1999	99US-0151103DP
PR	07-SEP-1999	99US-0151263DP
PR	10-SEP-1999	99US-0153707DP
PR	13-SEP-1999	99US-0153156DP
PR	15-SEP-1999	99US-0154041DP
PR	16-SEP-1999	99US-0154003DP
PR	20-SEP-1999	99US-0154177DP
PR	22-SEP-1999	99US-0155133DP

PR	23-SEP-1999;	99US-01554866;
PR	24-SEP-1999;	99US-01555699;
PR	26-SEP-1999;	99US-01564587;
PR	28-SEP-1999;	99US-01565967;
PR	04-OCT-1999;	99US-01571176;
PR	05-OCT-1999;	99US-01577533;
PR	06-OCT-1999;	99US-01578657;
PR	07-OCT-1999;	99US-01580259;
PR	08-OCT-1999;	99US-01582322P;
PR	12-OCT-1999;	99US-01583966;
PR	13-OCT-1999;	99US-01592932P;
PR	13-OCT-1999;	99US-01592949P;
PR	13-OCT-1999;	99US-01592952P;
PR	14-OCT-1999;	99US-01593929P;
PR	14-OCT-1999;	99US-01593930P;
PR	14-OCT-1999;	99US-01595931P;
PR	14-OCT-1999;	99US-01596337P;
PR	14-OCT-1999;	99US-01596348P;
PR	18-OCT-1999;	99US-01595848P;
PR	21-OCT-1999;	99US-0160741P;
PR	21-OCT-1999;	99US-01607614P;
PR	21-OCT-1999;	99US-01607688P;
PR	21-OCT-1999;	99US-01607707P;
PR	21-OCT-1999;	99US-0160814P;
PR	21-OCT-1999;	99US-0160815P;
PR	21-OCT-1999;	99US-01609801P;
PR	22-OCT-1999;	99US-01609812P;
PR	22-OCT-1999;	99US-01609898P;
PR	25-OCT-1999;	99US-01614040P;
PR	25-OCT-1999;	99US-01614054P;
PR	25-OCT-1999;	99US-01614066P;
PR	26-OCT-1999;	99US-0161639P;
PR	26-OCT-1999;	99US-01616360P;
PR	26-OCT-1999;	99US-01616516P;
PR	28-OCT-1999;	99US-01619200P;
PR	28-OCT-1999;	99US-01619932P;
PR	29-OCT-1999;	99US-0162142P;

Alignment Scores:	
Pred. No.:	1,13e-101
Score:	138.80
Percent Similarity:	52.84%
Best Local Similarity:	43.27%
Query Matchn:	37.92%
Ds:	3
Gaps:	21
Length:	166
Matches:	312
Conservative:	69
Mismatches:	116
Indels:	224
Gaps:	21

```

Qy      1 MetLysAspArgIleAsnGlnSerAsnAsnAsnThrProAspGlyAsnAspHisAsnThrPheGly 20
Db      1 ATGAAGTCTTTTGTGTATATGATGATATATATCATATGACACACACTATATTGTATGG 60

Qy      21 PheSerLeuSerProHis--MetLysMet----- 29
Db      61 TTCTCATGTGTTCTCAATATGATGAAAATGGAGGATGAGATGAGAACTATTAC 120

Qy      30 -----GluAlaThrSerAlaAlaIleValProThrThiPheMet 43
Db      121 TCATCTTCACTTCTTCAGCTGCACTTCTTCTTCTTCTGTTCCACCTCACTGTTGTT 180

Qy      44 SerProSerGlnSerHisLeuSerAspAspMetCysTyrGlyValGlyGluAspGly 63
Db      181 GGT-----GACAAACACTAGCAACTTTGGTGTGTGATAGATCTAACCCAAATGGA 231

Qy      64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerIleuCysTleLeu 83
Db      232 GGAATCTATTTCTACATGATGTGTGATGCCACTCAGATCTGATGGTTCTCTTGTGCTAATG 291

Qy      84 GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysIleuLys 102
Db      292 GAAAGTCTCAACAGATCTTCTCACTGCAATCACCATCAAGATTCATCTTCACAAAGGTGAG 351

```

QY	104	AspPheLeuGlyValAlaThrMetGlyThrHisGluTyrGlySerHisGlyValGly---	122
Db	352	GATTTCTTTGGG-----ACCATCAACAACAACAACAGTCAACAAGAACCATG	399
QY	123	---LeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnIleGluIleGlnProAsnArg	141
Db	400	GATCTTAGCTTAGATAGTATTATCTTACAAACACCTCAT-----GAGCCCAACAG	450
QY	142	AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln	157
Db	451	ACTACAACTTTCACAGAGTCTTTAGCTTCCCTCAACCAACAAACCAT-----	498
QY	158	ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGln	177
Db	499	-----GAGGAA	504
QY	178	GluThrThrIlyGlyGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu	197
Db	505	GAAACT-----AGAAATTACGGGAATGAC-----CCTAGTTTACACAT	543
QY	198	GlyLeuIlyAsnTprValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu	217
Db	544	GGA-----	546
QY	218	GlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly	237
Db	547	GGGTCTTTAATGATGAGGGGTATATGGGGAATTCAA-----	582
QY	238	CysGlyGlyIleLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal	257
Db	583	-----CACTCACTAGCTTATCCATGAGCCTGGCTGACAACTTCACTGCTCATC	630
QY	258	ThrAlaProSer-----	261
Db	631	ACTGGCTTCACACCAACCAACAAACAAACAAACCAACCAACCAACCAACG	690
QY	262	-----GlyThrAspSerValAlaVal	268
Db	691	CAGATCTCTGAAGCTCTTGAGAGCAAAAGCTTGGGTTTGAAGCAGCAGCATGGCGCT	750
QY	269	AspAlaIlyLysArgGlyHis-----AlaIlyLeuGlyGlnIlyGlnProValHis	285
Db	751	CGGAAGAGAGAGGGGCAAGAGAGTCTTGAAGTGTGGTCAGAAACAGATTGTCAT	810
QY	286	ArgIlySerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis	305
Db	811	AGAAATCATCATGACATCTTTGGACAAAGAACTTCTCAATCCGAGGGGTTCACAGACAT	870
QY	306	ArgTprThrGlyArgTyrGlyAlaHisLeuTprAspAsnSerCysIlyLeuGlyGlnGln	325
Db	871	AGATGACTGTGATATGATAGACTCATTTATGGGACATATGTTTCAGAAAGGAAGTCTAC	930
QY	326	ThrArgIlyGlyArgGlnValTyrLeuGlyGlyTyrAspMetGluGlnLysAlaAlaArg	345
Db	931	AGTAAAGAAAGCAAGCAAGTTATCTGGAGAGTTATGATATGAGAGAAAGCTCTCGA	990
QY	346	AlaTyrAspLeuAlaAlaLeuIlyTyrThrGlyProSerThrHisIleAspPheSerIle	365
Db	991	GGATATGATCTTGCTGCACTCAAGTACAGGGGTCCCTTCACTCACACCAATTTCTCTCG	1050
QY	366	GluAsnTyrGlnValGlnLeuGlnIleMetIlyAsnMetSerArgGlnGluTyrValAla	385
Db	1051	GAGAAATTATCAAGAAAGAGATTGAACATGAAAGAAATGACATGACAAAGAAATATGTGCA	1110
QY	386	HisLeuArgArgIlySerSerIlyPheSerArgGlyAlaSerIleTyrArgGlyValThr	405
Db	1111	CAATTGAGAGAGAGAGCACTGCTTTCTTAGGGGTGCTTCATCTTATAGAGGATCTCA	1170
QY	406	ArgHisHisGlnHisGlyArgTprGlnIleArgGlyLeuArgValAlaGlyAsnIlyAsp	425
Db	1171	AGACATACACACAGAGAAAGTGGCAAGCAGATTTGGATGAGTGGCTGGAAACAAACAT	1230
QY	426	LeuTyrLeuGlyThrPheSerThrGlnGluGluAlaIleGluAlaTyrAspValAlaIle	445

```

Db      1231 CTTACCTTGAACCTTTGGAAACCCAAAGAGTGGAGAGCTTACGATGTAGCAGCA 1290
Qy      446 ILeIysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGlu 465
Db      1291 ATTAAGTCCGGCGCAAAATGCTGTGACTAATTGATACACAGAGTACGAGTGTGAT 1350
Qy      466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgGlyIleAspAsn 485
Db      1351 CGTATCATGCTCTAGTAAACACTCTTGTCTGAGAGGTAGCGCAAGAAC----- 1401
Qy      486 AspProArgAsnIleAspIleAspTyrAsnIleAspValValThrSerValAsnAsnGlu 505
Db      1402 -----AACAAACAGCATTTGTC-----GTCAGGAAATACT 1428
Qy      506 GluThrValGlnValGlnAlaGlyAsnAsnAsnGluAsnAspSerGluTyrIleMet 525
Db      1429 GAA----- 1431
Qy      526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnIlys 545
Db      1432 -----GACCAAAAC 1440
Qy      546 IleMetAsnGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565
Db      1441 GCTCTAAATGCT----- 1452
Qy      566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerIleGly 585
Db      1452 ----- 1452
Qy      586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluAlaSer 605
Db      1453 -----GTTTGGAAGGTGTTCCAAACAAGAACTCAGT 1485
Qy      606 ProGluValArgGlyProSerLeu-----LeuPheProMetProProMetGluThrIlys 623
Db      1486 ACTCCCGAGAGACTCTTGAGTTTCCGGGAGATTTCGCGTTCCTCCCAAGTTAATCAAAAG 1545
Qy      624 IleVal-----AsnProIleGlyThrSerValThrSerTyrLeuProSerProThrValGln 642
Db      1546 ATGTTGCGATCAAAATATGCGCGGAATATAGAGTCCCTTGACATCAAAACCTTAATGCTAG 1605
Qy      643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrThrAsp 662
Db      1606 CTTAAG--ACCGTGCCTTACTTTGCTCTCAAGATCCGCTTTGCTGCTGCGCTGAT 1662
Qy      663 Thr 663
Db      1663 TCT 1665

```

```

XX      13-JUL-2000.
PD      07-JAN-2000; 2000MO-US0000465.
XX      PF
XX      08-JAN-1999; 99US-00227421.
XX      FR
XX      PA (REGC ) UNIV CALIFORNIA.
XX      PI Fischer RL, Mizukami Y;
XX      DR WPI: 2000-465969/40.
XX      DR P-FSDB: AAB07724.
XX      PT Modulating growth and cell proliferation in a plant used to alter organ
XX      PT mass, control fertility and enhance asexual reproduction in plants
XX      PT comprises modulating ANT activity and selecting plants with altered cell
XX      PT number.
PS      Claim 9; Page 40-41; 54pp; English.
XX      CC The present sequence encodes an aintegumenta (ANT) polypeptide. The ANT
XX      CC gene is expressed and functions not only in developing ovules but also in
XX      CC various developing organs. Growth and cell proliferation in plant can be
XX      CC modulated by modulating ANT activity. Modulation of ANT activity is used
XX      CC to alter organ mass, control fertility and enhance asexual reproduction
XX      CC in plants. Increased ANT activity can be used to produce male or female
XX      CC sterile plants. Inhibition of ANT activity can be used to truncate
XX      CC vegetative growth, resulting in early flowering
XX      SQ Sequence 2148 BP; 659 A; 431 C; 460 G; 598 T; 0 U; 0 Other;
SQ      Alignment Scores:
Pred. No.: 1 62e-101 Length: 2148
Score: 1328.00 Matches: 312
Percent Similarity: 52.84% Conservative: 69
Best Local Similarity: 43.27% Mismatches: 116
Query Match: 37.92% Indels: 224
DB: 3 Gaps: 21
US-10-024-632-2 (1-663) x AAAS9220 (1-2148)
Qy      1 MetIysArgIleAsnGluSerAsnAsnThrAspAspGlyAsnAsnHisAsnTyrPleuGly 20
Db      269 ATGAAGCTTTTGTGATATGATGATTAATCAATAGCAACACGACTAATTTGTTAGGG 328
Qy      21 PheSerLeuSerProHis--MetIysMet----- 29
Db      329 TTCTCATTTGCTTCAATATGATGATGAATGGAGGTAGAGAGTAGAGAGCTATTAC 388
Qy      30 -----GluAlaThrSerAlaIleThrValProThrThrPheTyrMet 43
Db      389 TCATCTTCAACTTCTTACGCTGCAACTCTTCTTCTTCTTCTTCACTCACTTGTGTT 448
Qy      44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGlyIleAsnGly 63
Db      449 GGT-----GACACACTAGCACTTTGGTGTGTTGCTATGAGATCAACCAATGGA 439
Qy      64 AsnPheHisSerProLeuThrValMetProLeuIleSerAspGlySerLeuCysIleLeu 83
Db      500 GAATCTATTTCTACAGTCTGTGATGCCACTGATCTGATGTTCTCTTCTTAATG 559
Qy      84 GluAlaLeuIleAsnArgSerGlnThrGlnValMetValProThrSerSerProIleLeuGlu 103
Db      560 GAAGCTCTCAACAGATTTTCTCACTGCAATCAACATCAATTCATTCCTCAAGGTGAG 619
Qy      104 AspPheLeuGlyGlyAlaThrMetGlyThrHisGluTyrGlySerHisGluArgGly-- 122
Db      620 GATTTCTTTGGG-----ACCATCAACAACAACAAGTCAACAAGAAGCATG 667
Qy      123 ---LeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArg 141
Db      668 GATCTTAGCTTAGTATTATTATTCTTACACACCACTCAT-----GAGCCCAACAG 718

```



QY 142 AspleuleuSerGlnProPhe-----ArgGlnGlnGlnHisMetSerValGln 157  
 Db 719 ACTACAACTTTCAAGAGTTCTTAGCTTCCCTCAACCGAAGAACAT----- 766  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGln 177  
 Db 767 -----GAGGAA 772  
 QY 178 GluThrThrLysGluThrHisValSerProCysSerSerLeuMetProGlnMetThrGln 197  
 Db 773 GAAACT-----AGAAATTACGGGAATGAC-----CCTAGTTGACACAT 811  
 QY 198 GlyLeuLysAsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGln 217  
 Db 812 GGA----- 814  
 QY 218 GlnGlnMetAsnCysGlyMetClyAsnGlnArgGlnGlyValSerLeuGlySerValGly 237  
 Db 815 GGGCTTTTAAATGATGGGGATATGGGGAAATTCA----- 850  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 Db 851 -----CAGTCACTGAGCTTATTCATGAGCCCTGGGTCACAACTTACCTGCATC 898  
 QY 258 ThrAlaProSer----- 261  
 Db 899 ACTGGCTTCACACACCAACCAAAACCAAAACCAACCAAGCCAAACCAACACAG 958  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 Db 959 CAGATCTCGAAGCTCTTGTCGAGCAAGCCCTGGGTTTGAGACGACACATGGCCGT 1018  
 QY 269 AspAlaLysLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285  
 Db 1019 GGGAGAAAGAAAGGGGGAACAAGAGATGTGTGATGTGGCAGAAACAGATTGTCAT 1078  
 QY 286 ArgLysSerIleAspThrThrGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
 Db 1079 AGAAATCTATCGATACCTTTGACACAGCACTTCCAAATCCGAGCGGTACAGACAT 1138  
 QY 306 ArgTrpThrGlyArgTyrGlnAlaHisLeuTrpAspAsnSerCysLysLeuGlyGln 325  
 Db 1139 AGATGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198  
 QY 326 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaArg 345  
 Db 1199 AGTAAAGAAAGAAAGCAAGATTTATCTGGAGGCTTGTGATGAGAGAGAAAGCTCTGA 1258  
 QY 346 AlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyProSerThrHisIleAsnPheSerIle 365  
 Db 1259 GCATATGATCTTGCTGCACTCAAGTACTGGGCTCCCTACTACACCAATTTCTCTCG 1318  
 QY 366 GluAsnTyrGlnValGlnLeuGlnLysMetLysAsnMetSerArgGlnGlnTyrValAla 385  
 Db 1319 GAAATTTATCAGAAAGATTGAAAGCATGAAAGCAATGATCAGCAAGAAATATGTTCA 1378  
 QY 386 HisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 405  
 Db 1379 CATTTGAG 1438  
 QY 406 ArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 425  
 Db 1439 AGACATCACACAGATGAAAGTGGCAAGCAGATGTTAGAGTGGCTGGAAACAAAGAT 1498  
 QY 426 LeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaGlnAlaGlyTyrAspValAla 445  
 Db 1499 CTTTACCTTGAACTTTTGAAACCAAGAAAGAGCTGCAAGAGCTTACATGATGACAGA 1558  
 QY 446 IleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGln 465  
 Db 1559 ATTAAGTTCCTGGCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1618

QY 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGlnLeuValAlaArgArgLysLysAspAsn 485  
 Db 1619 CGATCATGTCATGATGACACTCTTGTCTGAGAGTTAGCCGGAAGAAC----- 1669  
 QY 486 AspProArgAsnLysAspIleAspTyrAsnLysSerValAlaThrSerValAsnGln 505  
 Db 1670 -----AACCAACACATGTC-----CTCAGAACTACT 1696  
 QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnGlnLysAsnAspSerGlnTrpLysMet 525  
 Db 1697 GAA----- 1699  
 QY 526 ValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
 Db 1700 -----GACCAAAACC 1708  
 QY 546 IleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 Db 1709 GCTCTAAATGCT----- 1720  
 QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerLysIleGly 585  
 Db 1720 ----- 1720  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGlnAlaSer 605  
 Db 1721 -----CTTGGGAAGCGTCTCCACAAAGAGACAT 1753  
 QY 606 ProGlnLysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrLys 623  
 Db 1754 ACTCCGAGAGACTTGTGATTTCCGCGCATTTTGCCTGCTCCATGATATCAAAAG 1813  
 QY 624 IleVal---AsnProIleGlyThrSerValThrSerTrpLeuProSerProThrValGln 642  
 Db 1814 ATTTGGATCAAAATATGGCGGAATATGAGCTTGGACATCAAACTTAATGCTGAG 1873  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp 662  
 Db 1874 CTTAAG---ACCGTCCTCTTACTTGTGCTCAGATCCCGGTTTCCCTGCTTGGCGCTGAT 1930  
 QY 663 Thr 663  
 Db 1931 TCT 1933  
 RESULT 7  
 ID ABX13437 standard; DNA; 2148 BP.  
 XX ABX13437;  
 AC  
 XX 04-JUN-2003 (first entry)  
 DT  
 XX  
 DE A. thaliana AINTEGUMENTA DNA corresponding to Genbank U40256.  
 XX  
 KW Expression cassette; transgenic; promoter; LOX5; plant; food production;  
 KW animal feed; seed; stress resistance; disease resistance; starch content;  
 KW lipid content; dormancy; fibre content; pharmaceutical production;  
 KW fine chemical production; sterile plant; vitamin; flavouring; perfume;  
 KW dye; cotyledon; embryonic tissues; stress factor; LOX; ds.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN DE10127882-A1.  
 XX  
 PD 12-DEC-2002.  
 XX  
 PF 11-JUN-2001; 2001DE-01027882.  
 XX  
 BR 11-JUN-2001; 2001DE-01027882.  
 XX  
 PA (BADT ) BASF PLANT SCI GMBH.  
 XX  
 PI Bischoff F, Feussner I, Loyall LP;

XX WPI: 2003-279966/28.

XX Cassette for expressing transgene, useful e.g. in production of  
 PT pharmaceuticals and fine chemicals, contains promoter from the LOX5 gene  
 PT of Arabidopsis, provides cotyledon-specific expression.

PS Claim 5, Page: 28pp; German.

XX This invention describes a novel cassette for the transgenic expression  
 CC of nucleic acid containing the promoter of the LOX5 gene from Arabidopsis  
 CC thaliana or deletion variants of the LOX5 promoter which are functionally  
 CC linked to the nucleic acid of the invention. The cassette is used to  
 CC prepare transgenic organisms, especially plants, for production of foods,  
 CC animal feeds, seeds (including those with increased resistance to stress  
 CC and disease, altered starch/lipid contents or dormancy, or altered fibre  
 CC content), pharmaceuticals (especially antibodies, vaccines, enzymes and  
 CC pharmaceutical proteins) and fine chemicals (especially enzymes,  
 CC vitamins, amino acids, sugars, (un)saturated fatty acids, flavonoids,  
 CC perfumes and dyes), also to produce sterile plants. The LOX5 promoter  
 CC provides strong and specific expression in cotyledons and/or other early  
 CC embryonic tissue, so can degrade, or protect against, stress factors to  
 CC which these tissues are particularly sensitive. Since cotyledons are the  
 CC main storage organs of seeds, expressing transgenes in them produces  
 CC targeted increases/modifications in nutritional value. Expression in the  
 CC cotyledons is homogeneous, there are no side effects on other plant  
 CC organs (pollen) and the promoter is functional in a wide variety of  
 CC plants (ornamentals or crops). This sequence represents a nucleic acid  
 CC sequence associated with the Arabidopsis thaliana LOX gene described in  
 CC the disclosure of the invention

XX Sequence 2148 BP; 659 A; 431 C; 460 G; 598 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 1.62e-101 Length: 2148  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.92% Indels: 224  
 DB: Gaps: 21

US-10-024-632-2 (1-663) X ABX13437 (1-2148)

QY 1 MetLVsArgLIeAsnGLeuSerAsnAsnThrAspAspGlyAsnAsnH:SaenTryLeuGly 20  
 DB 269 ATGAAGCTCTTTTGGATATGATGATATATATCATAGCAACAGACTAATTTGTTGGG 328  
 QY 21 PheSerLeuSerProHis--MetLysMet----->----- 29  
 DB 329 TTCTCATGTGCTTCAATATGATGATAAATGGAGTAGAGAGAGTAGAGAGACTATTATAC 388  
 QY 30 -----GlualaHisSerAlaIaIaThrValProThrThrPheTyMet 43  
 DB 389 TCATCTTCACTTCTTCAAGCTGCAACTTCTTCTTCTTCTTCTTCACTTCACTTCTTCTT 448  
 QY 44 SerProSerGlnSerHisLeuSerAsnBheGlyMetCysTryGlyValGlyGluAsnGly 63  
 DB 449 GGT-----GACAACTACAGCAACTTGTGTTGCTATGATCTAACCCAAATGGA 499  
 QY 64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysAlaLeu 83  
 DB 500 GGAATCTATCTTCACATCTCTGTGATGCGCACTCGATCTGATGGTTCTTCTTGGCTTAATG 559  
 QY 84 GlualaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGlu 103  
 DB 560 GAAGCTCTCAACAGATCTTCTCATCTGCAATCACCATCAAGATTCATCTCCAAAGGTGGAG 619  
 QY 104 AspPheLeuGlyGlyAlaIaIaThrMetGlyThrHisGlyUryGlySerHisGluArgGly--- 122  
 DB 620 GATTTCCTTTGGG-----ACCATCTCAACAAACACAGCTACAAAGAAAGCCATG 667  
 QY 123 ---LeuSerLeuAspSerIleTryTyranSerGlnAsnAlaGluAlaGlnProAsnArg 141

DB 668 GATCTTAGCTTAGATAGTATTATCTACAAACCACTCAT-----GAGCCCAACAGC 718  
 QY 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln 157  
 DB 719 ACTAACAACCTTCAAGAGATTCTTTAGCTTCCTCCCTCAACCAAGAAACCT----- 766  
 QY 158 ThrHisProTryTyranSerGlyLeuAlaCysHisGlyLeuTyrglnAlaProLeuGluGlu 177  
 DB 767 -----GAGGAA 772  
 QY 178 GluThrThrLysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu 197  
 DB 773 GAAACT-----AGAAATTAAGGAAATGAC-----CCTAGTTTACACAT 811  
 QY 198 GlyLeuLysAsnThrPValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu 217  
 DB 812 GGA----- 814  
 QY 218 GlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 237  
 DB 815 GGGCTTTTAAATGATGAGGGGTATATGGGAAATTCAA----- 850  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 DB 851 -----CACTCACTGAGCTTATCCATGAGCCCTGGGTCAATCTAGCTGATC 898  
 QY 258 ThrAlaProSer----- 261  
 DB 899 ACTGGCTCTCACACCACCAACCAAAACCAAAACCAAAACCAAAAGCCAAACCCACAG 958  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 DB 959 CAGATCTCGAAGCTCTTGTGGAGACAAGCGTTGAGTGAAGACGACAGCAATGGCGCT 1018  
 QY 269 AspAlaLysLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285  
 DB 1019 GCGAAGAAAGAAAGGGGACAAAGAGAGTTGATGTTGTTGTCAGAAACAAATGTTTAT 1078  
 QY 286 ArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyranSerGlyValThrArgHis 305  
 DB 1079 AGAAATCATGATGATCTTTTGGACACAGAACTTCAATACCAAGGGCTTACAAAGCAT 1138  
 QY 306 ArgTryThrGlyArgTyrglnAlaHisLeuThrAspAsnSerCysLysLysGluGln 325  
 DB 1139 AGATGAGCTGATGATATGAGCTCATCTATGGCAATAGTTTCAGAGAGAGAGTAC 1198  
 QY 326 ThrArgLysGlyArgGlnValTyrglnGlyGlyTyranSerMetGluGluLysAlaIaArg 345  
 DB 1199 AGTAAAGAAAGAAAGCAAGTTATCTGGAGGTATGATATGAGAGAAAGAACTCTGCA 1258  
 QY 346 AlaTyranSerLeuAlaIaLeuLysTyrglyProSerThrHisIleAsnPheSerIle 365  
 DB 1259 GCATATGATCTTGTGCACTCAAGTACGGGTCCCTCTACTACACCAATTTCTCTGGC 1318  
 QY 366 GluAsnTyrglnValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyranAla 385  
 DB 1319 GAGAAATTAACAAGAAAGATTGAAGACATGAAGAAACATGATCAAGCAATATTTGCA 1378  
 QY 386 HisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyranArgGlyValThr 405  
 DB 1379 CATTTGAGAGGAAGAGAGAGCTGTTCTCTAGGGGCTTCATCATATGAGAGATGACA 1438  
 QY 406 ArgHisHisGlnHisGlyArgTyrglnAlaArgIleGlyArgValAlaGlyAsnLysAsp 425  
 DB 1439 AGACATCACACACATGAGAGAGGAGACGAGATTGGTAGTGCTGGAAACAAAGAT 1498  
 QY 426 LeuTyranGlyTyranPheSerThrGlnGlnGluAlaIaGluAlaTyranValAlaIa 445  
 DB 1499 CTCTACCTTGGAACTTTGGAAACCAAGAAAGAGCTGCAAGGCTTAGCATATACAGCA 1558  
 QY 446 IleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyranValGlu 465  
 DB 1559 ATTAAGTTCCGTGGCAAAATGCTGTGACTAATTGATATACGAGGATGATGATGAT 1618

```

QY 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgIleSerAspAsn 485
XX 1619 CGATTCATGCTTACTTAACACACTCTTGTCTGAGAGATTACCGCAAGAAC----- 1669
QY 486 AspProArgAsnIleAspIleAspIleAspIleAspIleAspIleAspIleAspIle 505
XX 1670 -----AACAAACGCAATGTC-----GTCAGCAATACT 1696
QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnGluAsnAspSerGluTrpIleMet 525
XX 1697 GAA----- 1699
QY 526 ValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlySerAspGlnIle 545
XX 1700 -----GACCAAAAC 1708
QY 546 IleMetAsnGlyGlyAsnTrpArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565
XX 1709 GCTTAAATGCT----- 1720
QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerIleGly 585
XX 1720 ----- 1720
QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluAlaSer 605
XX 1721 -----GTTGTGAAAGCTGTTCCAAAGAAAGTCACT 1753
QY 606 ProGluArgGlyProSerLeu-----LeuPheProMetProPometGluThrIle 623
XX 1754 ACTCCCGAGAGACTCTTGAGTTTCCGCGCATTTCCGCTGCTCAAGTAAATCAAAAG 1813
QY 624 IleVal-----AsnProIleGlyThrSerValThrSerIleProSerProThrValGln 642
XX 1814 ATGTCGGATCAATATATGCGAATATGAGTCTTGACATCAAACTTAATGCTGAG 1873
QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp 662
XX 1874 CTTAAG---ACCGTCGCTTACTTGGCTCAGATGCCGTTTCCGCTGCTGGGCTGAT 1930
QY 663 Thr 663
XX 1931 TCT 1933
Db 1931 TCT 1933

RESULT 8
ID AAA59222 standard: cDNA: 1738 BP.
XX AAA59222;
AC AAA59222;
XX 07-NOV-2000 (first entry)
DE cDNA encoding a Brassica aintegumenta (ANT) polypeptide.
XX
XX Aintegumenta; ANT; cell proliferation; growth; organ mass; fertility;
XX asexual reproduction; plant; male sterile plant; female sterile plant;
XX early flowering; ss.
XX
XX Brassica sp.
OS
XX
XX Key location/Qualifiers
XX CDS 1..1647
XX /*tag= a
XX /product= "aintegumenta polypeptide"
XX
XX MO200040694-A2.
XX
XX 13-JUL-2000.
XX
XX 07-JAN-2000; 2000MO-US000465.
XX
XX 08-JAN-1999; 99US-00227421.

```

```

XX (REBC ) UNIV CALIFORNIA.
PA Fischer RL, Mizukami Y,
XX MPI: 2000-465969/40.
XX P-PSDB; AAB07725.
XX
XX DR Modulating growth and cell proliferation in a plant used to alter organ
XX PT mass, control fertility and enhance asexual reproduction in plants
XX PT comprises modulating ANT activity and selecting plants with altered cell
XX PT number.
XX
XX PS Claim 87, Page 43-44; 54pp; English.
XX
XX CC The present sequence encodes an aintegumenta (ANT) polypeptide. The ANT
XX CC gene is expressed and functions not only in developing ovaries but also in
XX CC various developing organs. Growth and cell proliferation in plant can be
XX CC modulated by modulating ANT activity. Modulation of ANT activity is used
XX CC to alter organ mass, control fertility and enhance asexual reproduction
XX CC in plants. Increased ANT activity can be used to produce male or female
XX CC sterile plants. Inhibition of ANT activity can be used to truncate
XX CC vegetative growth, resulting in early flowering.
XX
XX SQ Sequence 1738 BP; 500 A; 361 C; 412 G; 465 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 6,996-100 Length: 1738
XX Score: 1307.00 Matches: 308
XX Percent Similarity: 54.52% Conservative: 78
XX Best Local Similarity: 45.50% Mismatches: 104
XX Query Match: 37.32% Indels: 218
XX DB: Gaps: 25
XX
XX US-10-024-632-2 (1-663) x AAA59222 (1-1738)
XX
XX QY 9 AsnThrAspAspGlyAsnAsnHisAsnTrpLeuGlyPheSerLeuSerProHisMet--- 27
XX 16 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 75
XX 28 -----LysMetGluAlaThrSerAlaAlaThrValProThrPhe 41
XX 76 AAAATGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132
XX QY 42 TyrMetSerProSerGln-----SerHisLeuSerAsnMetGlyMetCysIle 57
XX 133 ---TGTGTTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 189
XX QY 58 GlyValGly---GluAsnGlyAsnPheHisSerProLeuThrValMetProLeuIleSer 76
XX 190 GGTTCCTAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 249
XX QY 77 AspGlySerLeuCysIleLeuGluAlaLeuIleValSer----- 89
XX 250 GACGGTTCCTTTGCTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
XX QY 90 GlnThrGluValMetValProThrSerSerProIleGluAspPheLeuGlyGlyAla 109
XX 310 CATAGTCAAGTT-----TCATTCCAAGATGAGAGATTC----- 345
XX QY 110 ThrMetGlyThrHisGluTrpGly---SerHisGluValGly-----LeuSerLeuAsp 126
XX 346 ---TTTGGAGCCATCATCAACAAAGTCAAAAGCCATGATCTTGTAGTATGAT 402
XX QY 127 SerIleTrpTrpAsnSerGlnAsnAlaGlnAlaGlnProAsnArgAspLeuSerGln 146
XX 403 AGTTTATTTCTACAAATACACTCATGCG-----CCAAACAAACACCAACTTTCAA 453
XX QY 147 ProPheArgGlnGlnGlnIleHisMetSerValGlnThrHisProIleTrpSerGlyLeuAla 166
XX 454 GAGTTC-----TTAGCTTCCTCAAACTGAAACACCAT----- 489
XX QY 167 CysHisGlyLeuTrpGlnAlaProLeuGluGluGluThrThrIleGluThrHisValSer 186

```





Db 1522 -AATGCCAACAACAGTATGATGACACACAGAT----- 1554  
 Qy 557 eserMetAlaLeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMe 577  
 Db 1554 ----- 1554  
 Qy 577 tLeuAspGlySerSerIleGlyIleHisPheSerAsnThrSerSerLeuValThre 597  
 Db 1555 -----GAGAGTGT-----GATGCTGAATTCACACACACAT-----GTAACCTGA 1592  
 Qy 597 rLeuSerSerSerArgGlnAlaSerProGlnulysArgGlyProSerSerLeuPheProMe 617  
 Db 1593 ATCATCATCATTCACACACAGTTTCACACACAGC-----AACACAGAGCTTTGACACCTCA 1649  
 Qy 617 t-ProprometGluThrIleValAsnProIleGlyThrSerValIleThrSerIleuP 637  
 Db 1650 AAGCCCAATGAGTTTGAGTCAAGTGGAGCAGATTATGACATGGCTACTTCAACCTTGA 1709  
 Qy 637 roSerProThrValGlnMet-----ArgProSerProAlaI 649  
 Db 1710 TGGACCAAAATATATATATGAGGAAACATGAAATGACCAATGACACACAGAGCTTGG 1769  
 Qy 649 leSerLeu-----SerHisLeuProValPheAlaSerIleThrAsp 662  
 Db 1770 A-AATTGGGATGGTTAATCAAGTTCACATGTTGCTTTGTGGAATGAA 1818  
 RESULT 10  
 ACC00803  
 ID ACC00803 standard; cDNA; 2463 BP.  
 AC ACC00803;  
 XX 16-MAY-2003 (first entry)  
 DT Glycine max oil trait related cDNA sequence SEQ ID NO:355.  
 DE  
 XX Glycine max oil trait related cDNA sequence SEQ ID NO:355.  
 XX  
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 XX receptor-like protein kinase; mitogen activated protein kinase; oil;  
 XX LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 XX CAC-like transcription factor; antisense inhibition; co-suppression;  
 XX transgenic plant; gene; ss.  
 OS Glycine max.  
 XX  
 XX WO2003002751-A2.  
 XX  
 XX 09-JAN-2003.  
 XX  
 XX 27-JUN-2002; 2002MO-US020152.  
 XX  
 XX 29-JUN-2001; 2001US-0301913P.  
 XX  
 XX (DUBO) DU PONT DE NEMOURS & CO E I.  
 XX (PION-) PIONEER HI-BRED INT INC.  
 XX  
 XX Allen SW, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT,  
 XX Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 XX Tarczynski MC;  
 XX  
 XX WPI; 2003-201509/19.  
 XX P-PDB; ABR40768.  
 XX  
 XX Novel nucleotide fragment encoding polypeptides having receptor-like  
 XX protein kinase activity, caleosin-like activity, useful for altering oil  
 XX phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 XX Claim 18; Page 373-374; 542pp; English.  
 XX  
 XX The present invention describes an isolated nucleotide fragment (I)  
 XX comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 XX polypeptide (P) having receptor-like protein kinase activity, mitogen  
 XX activated protein (MAP)-kinase activity, LIP5-like transcription factor

CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity and KCC-like transcription factor activity. Also described (1) or  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (II) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to  
 CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention  
 SQ Sequence 2463 BP; 770 A; 539 C; 521 G; 633 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7,15e-79 Length: 2463  
 Score: 1059.50 Matches: 288  
 Percent Similarity: 47.59% Conservative: 77  
 Best Local Similarity: 37.55% Mismatches: 202  
 Query Match: 30.25% Indels: 201  
 DB: 7 Gaps: 23  
 US-10-024-632-2 (1-663) x ACC00803 (1-2463)  
 Qy 17 AsnThrLeuGlyPheSerLeuSerProHisMetIleValIleThrSer---AlaAla 35  
 Db 197 AATTGTTAGGTTTCA 256  
 Qy 36 ThrValProThrThrPheThrMetSerPro----- 45  
 Db 257 ACGGACCTCTCTGTTTGTCTTCAACCTTGATGAAATCTCAAGCATGATGATGACAGA 316  
 Qy 46 -----SerGlnSerHisLeuSerAsnPheGlyMetCysTyr 57  
 Db 317 GACTGCTTTCATCTCACTTCCTGATCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTAC 373  
 Qy 58 GlyValGlyGlu-----AsnGlyAsnProHisSerProLeuThrValMetPro 73  
 Db 374 GGCATATACGAGACCTTTTCATAGAGCAACATATTCAC-----ACCATCAAGAT 424  
 Qy 74 LeuIleSerAsp-----GlySerLeuCysIleLeuGluAla 85  
 Db 425 TGGAGAGACATACACACACCAAACTTGCTATTGGAACTTCATGACAGCAACCAAC 484  
 Qy 86 LeuIleArgSerGlnThrGlnValMetValProThrSerSerProIleLeuGlnAspPhe 105  
 Db 485 ATGAACCAACACCATCAGCAGCAACCA-----CAACACAGCCCAAGCTTGAAACCTTC 538  
 Qy 106 LeuGlyGlyAlaThrMetGlyThrHisGlyIleGlySerHisGlyArgGlyLeuSerLeu 125  
 Db 539 CTGGGTGGA-----CACTCATTTGGTGAACATGAG----- 568  
 Qy 126 AspSerIleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuSer 145  
 Db 568 ----- 568  
 Qy 146 GlnProPheArgGlnGlnGlnIleHisMetSerValGlnThr-----His 159  
 Db 569 CAACCTTAC-----GATGTAACCTCAGCTCTACAGAAATATACATGTTCCGGCTCAG 619  
 Qy 160 ProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlnGlnThr 179  
 Db 620 CCGGTATTTGGCGCGGTGGCGCGCGGTGGT----- 649  
 Qy 180 ThrIleGlnThrHisIleValSerAspCysSerSerLeuMetProGlnMetThrGlnGlyLeu 199  
 Db 650 AGCAATAGACAGCAACACACACAGTACTTCATCGGTTATCCATG-----ATA 700  
 Qy 200 LysAsnThrValAlaIleProThrArgGlnPheSerThrHisGlnGlnValLeuGlnGln 219  
 Db 701 AAGACATGTTG-----AGAACCAACGACCACTCAGAAACCAACATTAACAC 751

QY 220 MetAsnCyseGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGly 239  
 DB 752 AACATGATAAGGTGGCAATAGAGAAAGCATGTGCAG----- 790  
 QY 240 GluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAla 259  
 DB 791 -----CAGACTATACCTTCATGAGTACTGGTTCCACAAACAGC--ACATACATA 841  
 QY 260 ProSerGlyThrAspSerValAlaValAspAlaLysLysArgGlyLysAlaLysLeuGly 279  
 DB 842 CCCCTTCACCTGCTAGTGTG-----GATATGAGAGAGATTCTCTGAT 886  
 QY 280 GlnLysGlnPro----- 283  
 DB 887 AACAAAGAACCCATACACGCGCTGCTGATATACAAACCAACCGGAGCCATTGAAC 946  
 QY 284 ValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThr 303  
 DB 947 GCACCCAGAAAGTCCATTGACATTTGGACAGAGAACTTCTATCAGCGTGGTATACA 1006  
 QY 304 ArgHisArgTyrPheGlyArgTyrGlyAlaHisLeuThrAspAsnSerCysLysLysGln 323  
 DB 1007 AGGATATGGTGGAGCGGGAGGTATGAGGCTCACCTGGGATATATGTTGTAGAGAGAG 1066  
 QY 324 GlnGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAla 343  
 DB 1067 GGACAACTCCGAAAGAGAGAGAGAGTTTACTTGGAGAGTTATGACAAAGAGAGAGAGAG 1126  
 QY 344 AlaArgAlaTyrAspLeuAlaLysLysTyrTyrGlyProSerThrHisIleAsnPhe 363  
 DB 1127 GCTAGAGCTCATGATTGGACGACCTAAATATCGGGAAACAACTAGCAACAAATTTT 1186  
 QY 364 SerIleGluAsnTyrGlnValGlnLeuGlnGluMetLysAsnMetSerArgGlnGlnTyr 383  
 DB 1187 CCAATTACCCACTAGAGAAAGAGTGGAGAAATGAGACCAATGACTAGGCAAGATAC 1246  
 QY 384 ValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGly 403  
 DB 1247 GTTGGCTCATGAGAGAGAGAGAGTATGCGGTTTCTGCGGGGATCATTTATCGAGGT 1306  
 QY 404 ValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgTyrGlyValAlaGlyAsn 423  
 DB 1307 GTGACGAGACACCATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1366  
 QY 424 LysAspLeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaGlnAlaTyrAspVal 443  
 DB 1367 AAGGATCTCTACTTGGAACTTTCCGACCCAAAGAGAGAGAGAGAGAGAGAGAGAG 1426  
 QY 444 AlaAlaIleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAsp 463  
 DB 1427 GCAGCAATCAATTCAGAGAGAGAGAGAGTCTGTCAAACTTGCATGAGAGAGATGAC 1486  
 QY 464 ValGlnAlaGlnIleMetAlaSerSerAsnLeuAlaGlyLysValAsnArgLysLys 483  
 DB 1487 GTGAAAACATATCTGAGAGACACCTTGCATTTGGTGTCTGCAAAACCGTTTGAG 1546  
 QY 484 Asp-----AsnAspProArgAsnLysAspIleAspTyr 494  
 DB 1547 GATTATGAG 1606  
 QY 495 AsnLysSerValValThrSerValAsnAsnGlnGlnValThrValGlnValGlnAlaGlyAsn 514  
 DB 1607 ACTACACTATGAA-CTCTCACTTACTCAAGAACTCTTAACAACATTCAGAGAGAGAG 1665  
 QY 515 AsnAsnAsnLysAsnAspSer-GluTyrLysMetValLeuPheAsnHisProSerGlnGly 534  
 DB 1666 AACACACAGGAGACTCATCATCACTGAGACAAAGCTCTTGCATTCACCAACCTCAACC 1725  
 QY 534 ngInGlnAlaAsnGlyAsnGlySerAspGlnLysIleMetLysCysGlyAsnTyrArg- 553  
 DB 1726 TTGCACCACTATCACTACCTTATGACAAAGAAATTAATGTGTGCAAGCAAGACAGA 1785

QY 554 -----AsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyLysAspSerValGly 571  
 DB 1786 CAACTGTATGCTCTCTACTCTTTGCTCTTATTCAGATTTATATCACTACAGCTAGGGAA 1845  
 QY 571 YserGlyGlnHisAsnMetLeuAspGluSerSerLysIleGlyThrHisPheSerAsnTh 591  
 DB 1846 CAATGGACACACAACTTCTTACACAAATTCAGAGGTGGACCCCTATGTTTAAACATGCA 1905  
 QY 591 rSerSerLeuValThrSerLeuSerSerSerArgGlnAlaSerProGluLysArgGlyPro 611  
 DB 1906 TTCTGCTTCATGACAACTACTCTTCACTCACTCTGTTTATATGATGTATGAGAG 1965  
 QY 611 cSerLeuLeuPheProMetProPrometGluThrLysIleVal----- 625  
 DB 1966 TGTGGGGGCTATATATGATGATTCCTATGGGAGCTACTACTGTGTGCAATGATGG 2025  
 QY 626 -----AsnPro----- 627  
 DB 2026 TGATCAAAATCCAGAGAGCAATCATGCTTTGGTATATGATAGATAAGCACTTGTTA 2085  
 QY 627 ----- 627  
 DB 2086 TGAAGTGTATGTGTCTTACACTGATCCTTATCATGACATGCAAGAACTGTATTA 2145  
 QY 628 -----IleGly 629  
 DB 2146 TCTTACTCAACAGCAACATCTTCTGTATGATGACAGTGAAGCTATGATCAAGG 2205  
 QY 629 YThrSerValThrSerThrPheProSer-----ProThrValGlnMetArgProSerPro 647  
 DB 2206 ATCTGATGCAATATCTTGGTTCCTCACTGCTATTCCACTCATGACCAAGTCTAGTAC 2265  
 QY 647 cAlaIleSerLeuSerHis 653  
 DB 2266 TAGTATGGCTCTCTGCCAT 2284  
 DB  
 RESULT 11  
 ACC00801  
 ID ACC00801 standard; cDNA; 2588 BP.  
 AC ACC00801;  
 XX  
 DT 16-MAY-2003 (first entry)  
 XX  
 DE Glycine max oil trait related cDNA sequence SEQ ID NO:351.  
 XX  
 DE Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 KM receptor-like protein kinase; mitogen activated protein kinase; oil;  
 KM lipid-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 KM CKC-like transcription factor; antisense inhibition; co-suppression;  
 KM transgenic plant; gene; ss.  
 XX  
 OS Glycine max.  
 XX  
 PN W02003002751-A2.  
 XX  
 PD 09-JUN-2003.  
 XX  
 PF 27-JUN-2002; 2002WO-US020152.  
 XX  
 PR 29-JUN-2001; 2001US-0301913P.  
 XX  
 PA (DUPC) DU PONT DE NEMOURS & CO E I.  
 XX (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Farnedu OO, Harvell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczyński MC;  
 XX  
 DR WPI: 2003-201509/19.  
 DR P-PSDB; ABR40766.  
 XX  
 PT Novel nucleotide fragment encoding polypeptides having receptor-like













AC ACC00799;  
 XX 16-MAY-2003 (first entry)  
 XX Glycine max oil trait related cDNA sequence SEQ ID NO:347.  
 DE  
 XX plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 XX receptor-like protein kinase; mitogen activated protein kinase; oil;  
 XX lipid-like transcription factor caleosin; ATP citrate lyase; SMT;  
 XX CAC-like transcription factor; antisense inhibition; co-suppression;  
 XX transgenic plant; gene; ss.  
 OS Glycine max.  
 XX WC0003002751-A2.  
 XX 09-JAN-2003.  
 XX 27-JUN-2002; 2002WC-US020152.  
 XX 29-JUN-2001; 2001US-0301913P.  
 XX (DUP0 ) DU PONT DE NEMOURS & CO E I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX Allen SM, Allen WB, Cahoon RE, Bredaun S, Famodu OO, Harvell LT,  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B,  
 PI Tarczyński MC;  
 XX WPI; 2003-201509/19.  
 DR P-PSDB; ABR40764.  
 XX Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity, caleosin-like activity, useful for altering oil  
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX Claim 18; Page 364-365; 542pp; English.  
 PS The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
 CC activated protein (MAP)-kinase activity, lipid-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SMT-like  
 CC activity and CAC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV) or  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00826 to ACC00868 and ABR40591 to  
 CC ABR40893 represent sequences used in the exemplification of the present  
 CC invention  
 CC  
 XX Sequence 1874 BP; 542 A; 472 C; 412 G; 448 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 4.75e-70 Length: 1874  
 Score: 952.50 Matches: 248  
 Percent Similarity: 49.84% Conservative: 63  
 Best Local Similarity: 39.74% Mismatches: 153  
 Query Match: 27.20% Indels: 160  
 DB: Gaps: 19  
 US-10-024-632-2 (1-663) x ACC00799 (1-1874)  
 QY 17 AANTPLGGLYPheserLeuserProHismetLysMetGluAlaThrSeraAlaAlaThr 36  
 DB 184 AACTGCTTCCTCTCTCTCTCCCA--ATGGAATGCTCCGAACCTCCGAACCTGAC 240  
 QY 37 ValProThrTherPhetYrMetserProserGlnserHisLeuserAsnPhetGlyMetCys 56

DB 241 TTCTTCATATACAGCGCGCTTCGGTCTCTCTCATCTCTCACTACTACTGCAACTG 300  
 QY 57 TygGlyValGly---GluAsnGlyAsnPh-----HisSerPro 68  
 DB 301 TACACCAACGCGTGGGCAACGAGCTCAAGTTTGACAGATCTCAACCAACGAGTAC 360  
 QY 69 LeuThrValMetProLeuLysSerAspGlySerLeuCySileLeuGluAlaLeuLysArg 88  
 DB 361 GTGAGTTTGTGA-----TCTTCG 381  
 QY 89 SerGlnThrGlnValMetValProThrSerSerProLysLeuGluAsnPhLeuGlyGly 108  
 DB 382 TCGCAGACGCGCGCCAGACGAGACGAGACGAGCTCTCG-----CTGACGACATCTAC 441  
 QY 109 AlathMetGlyThrHisGlyTyGlySerHisGluArgGlyLeuSerLeuAspSerIle 128  
 DB 442 TCTTCGCGC-----GTGATG 456  
 QY 129 TyrTyraSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPhe 148  
 DB 457 CGTACTCTCCAGACCCAGACGAGACGAGACGAGCTCTCG-----CTGACGACATCTAC 510  
 QY 149 ArgGlnGlnGlyHisMetSerValGlnThrHisProTyTyTySerGlyLeuAlaCyHis 168  
 DB 511 GACCACCAACCAAC-----CACCAACCAACCAACCGCTTACTCTCTCTCTCTC 558  
 QY 169 GlyLeuTyGlnAlaProLeuGluGlu-----GluThrThrLys 181  
 DB 559 GGTGTGACACGACGAGATCTCAAGCCATTCTGATTTCAAGCTTTTCGACCAACTCC 618  
 QY 182 GluThrHisValSerAspCySerSerLeuMetProGlnMetThrGluGlyLeuLysAsn 201  
 DB 619 GGTTCGAGGTGATGATCTGCAATCCATC----- 648  
 QY 202 TrpValAlaProThrArgGluPheserThrHisGlnGlnValLeuGluGlnMetAsn 221  
 DB 649 GGAAGAAGCGACGCGAGAGTTCGAGCTCACTCT-----ATTGAGTCTCTCC----- 696  
 QY 222 CysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCyGlyGluLeu 241  
 DB 697 -----GGCAACGAGTTCGCGCGCTTCGCGTGCACACCGCA----- 735  
 QY 242 GlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaProSer 261  
 DB 736 ---ACCTTCGCTCGCGCTTGACATGAGCTCCGAGAAAGCGCTTCGCGCGAGATCC 792  
 QY 262 GlyThrAspSerValAlaValAspAlaLysLysArgGlyHisAlaLysLeuGlyGlnLys 281  
 DB 793 AATAGCTCG----- 801  
 QY 282 GlnProValHisArgLysSerIleAspThrPhetGlyGlnArgThrSerGlnTyArgGly 301  
 DB 802 -----AAGAATGCTGATATCTTCGCGCAGCGAGCTTCAATTACAGAGGT 849  
 QY 302 ValThrArgHisArgTyrThrHisGlyArgGlyGluAlaHisLeuThrAspAsnSerCys 321  
 DB 850 GTTACTAGGACACCGATGACAGAGAAATATGAGGCACTCTATGGACATATGTTGCA 909  
 QY 322 LysGluGlyGlnThrArgGlyValArgGlnValTyLeuGlyGlyTyAspMetGluGlu 341  
 DB 910 AGGAGGCTGACCGCAAGAAAGGCGTCAAGTTTATTTGGTGATGATGAAGAAGA 969  
 QY 342 LysAlaAlaArgAlaTyArgAspLeuAlaLeuLysTyTyTyGlyProSerThrHisIle 361  
 DB 970 AAGCGCGAGAGCTTATGATTTGCGAGCTTAAAGTACTGGGCTCTCCACTCTACACAC 1029  
 QY 362 AsnPhSerIleGluAsnTyGlnValGlnLeuGluGluMetLysAsnMetSerArgGln 381  
 DB 1030 AACTTCCTGTTTCATATTCGAAAGAGAGGAGAGATGAACATCTAACAAAGCA 1089  
 QY 382 GluTyValAlaHisLeuArgArgLysSerSerGlyPheserArgGlyAlaSerIleTy 401

```

Db      1090 GAATTTATTCATCATTCGGGAGAGAAAGTAGTGTTCCTCCAGGGAGCTTCATATAC 1149
Qy      402 ATGGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
Db      1150 AGAGGTGTTTCAAGGATCATCAACAGGATGATGATGATGATGATGATGATGATGATGAT 1209
Qy      422 GATATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 441
Db      1210 GGAACCAAGATTTATCTTGGGACATTCGACACCGAGAGAGACACAGAGGATAT 1269
Qy      442 AspValAlaAlaIleLeuPheAsnGlyAlaAsnAlaValThrAsnPheAspIleSerArg 461
Db      1270 GATATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
Qy      462 TyrAspValGluArgGluMetAlaSerSerAsnLeuAlaGluLeuAlaArg 481
Db      1330 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
Qy      482 LysLysAspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValValThrSer 501
Db      1387 TTGAGG----- 1392
Qy      502 ValAsnAsnGluGluThrValGlnValGlnAlaGlyAsnAsnAsnGluAsnAspSer 521
Db      1393 -----CTTCCCTTGAAATCA 1407
Qy      522 GluTyrPheMetValLeuPheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGly 541
Db      1408 GAGCGAAGAGCTCCTCCTGATGACAGACAGACAGATGACAGATGACAGATGACAGATGAC 1455
Qy      542 SerAspGlnLysIleMetAsnGlyAsnTyrArgAsnSerAla----- 556
Db      1456 -----CAGTGTGTGTAACGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1497
Qy      557 -PheSerMetAlaLeuGlnAsnLeuGlyIleAspSerValGlySerGlyGlnHisAs 576
Db      1498 CATCGCCCAATTCCTCAATCCCTTGTC-----GAAATCCG 1533
Qy      576 MetLeuAspGluSerSerLysIleGlyThrHisPheSerAsnThrSerSerLeuValTh 596
Db      1534 TTTCATTCACACACAGCATATTATTCCTACCAACCTTTCCCAACATTTTCAACCTTACCA 1593
Qy      596 rSerLeuSerSerSerArgLysAlaSerProGluLysArgGlyProSerLeuLeuPheP 616
Db      1594 GCTGTGTC-----CAGACGCTGCTGCTTACTTCTGCC 1626
Qy      616 QMetProPro 619
Db      1627 AATGCAACCG 1636

RESULT 15
ADD30878
ID      ADD30878 standard, cDNA, 1818 BP.
XX      ADD30878;
XX      15-JAN-2004 (first entry)
DE      Plant yield-related polynucleotide clone GI793.
KW      ds; transcription factor; transgenic plant; growth rate; senescence;
XX      seed germination rate; plant vigor; seedling vigor.
XX      Arabidopsis thaliana.
XX      OS
XX      PN      WO2003013227-A2.
XX      PD      20-FEB-2003.
XX      PF      09-AUG-2002; 2002WO-US025805.
XX      PR      09-AUG-2001; 2001US-0310847P.
XX      PR      19-NOV-2001; 2001US-0336049P.

```

```

PR      11-DEC-2001; 2001US-0338692P.
PR      14-JUN-2002; 2002US-00171468.
XX      (MEND-) MENDEL BIOTECHNOLOGY INC.
XX      Ratcliffe O, Riechmann UJ, Adam LJ, Dubell AT, Heard JE;
XX      Pilgrim MD, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
XX      Brown PE;
XX      WPI; 2003-248221/24.
XX      P-PEDB; ADD30879.
DR      11-DEC-2001; 2001US-0338692P.
PT      New plant transcription factor polynucleotides and polypeptides, useful
PT      in producing transgenic plants with commercially valuable properties,
PT      such as an alteration in a plant growth characteristic, e.g. growth rate
PT      or apomixis.
PS      Disclosure; SEQ ID NO 907; 454bp; English.
XX      CC      The invention relates to a number of isolated Arabidopsis thaliana cDNA
XX      CC      sequences and their encoded proteins which are especially transcription
XX      CC      factor related cDNA's and proteins. The isolated or recombinant plant
XX      CC      transcription factor polynucleotides and polypeptides are useful in
XX      CC      producing transgenic plants with commercially valuable properties, i.e.
XX      CC      modified or altered desirable traits as compared to a reference plant,
XX      CC      such as an alteration in a plant growth characteristic, e.g. growth rate,
XX      CC      germination rate of seeds, vigor of plants and seedlings, or leaf and
XX      CC      flower senescence. Sequence information related to the polynucleotides
XX      CC      and polypeptides can also be used in bioinformatic search methods. The
XX      CC      transgenic plant is useful for growing a progeny plant from a parent
XX      CC      plant. This sequence represents one of the cDNAs of the invention.
XX      SQ      Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,528-70 Length: 1818
Score: 951.50 Matches: 252
Percent Similarity: 48.61% Conservative: 80
Best Local Similarity: 36.90% Mismatches: 157
Query Match: 27.17% Indels: 194
DB: 9 Gaps: 23

US-10-024-632-2 (1-663) x ADD30878 (1-1818)
Qy      14 AsnAsnHisAsnTyrLeuGlyPheSerLeuSerProHisMetLysMetGluAlaThrSer 33
Db      62 AATCTTAACAACGCTGCTGCTTCTCCTTTCACCGAAC-----AAC 103
Qy      34 AlaAlaThrValProThrThrPheThrMetSerProSerGlnSerHisLeuSerAsn--- 52
Db      104 TCTTCTTTCCTCCTCCATCAATACCACTTGCTGTCAGGACCATATGACAAACCT 163
Qy      53 -----PheGlyMetCysTyrGlyValGlyGluAsnGlyAsnPheHisSer 67
Db      164 TTTCACACACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214
Qy      68 ProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeuGluAlaLeuLys 87
Db      215 -----GAAAGAGGA----- 223
Qy      88 ArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGluAspPheLeuGly 107
Db      224 -----GAGGTTCCAAAAGTGCGCATTTTCTCGGT 253
Qy      108 GlyAlaThrMetCysThrHisGlyGlySerHisGluArgGlyLeuSerLeuSerSer 127
Db      254 -----GTGAGCAACCGGACGAAACCAATCCACACCTTGATGACTTACACACACTA 307
Qy      128 IleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuSerGlnPro 147
Db      308 GACTACTACTTCCATCCCAATAGC---TTGATGCTCT----- 340
Qy      148 PheArgGlnGlnGlnHisMetSerValGlnThrHisProTyrTyrSerGlyLeuAlaCys 167

```

```

Db      341 -----AGCGTCCATCAAC----- 355
Qy      168 HisGlyLeuTyrglnAlaProLeuGluGluGluThrThrlsGluThrHisValSerAsp 187
Db      356 -----GATCGCTGTGAGCAGCT 373
Qy      188 CysSerSerLeuMetProGlnMetThrGluGlyLeuLeuSerAsnTrpValAlaProThrArg 207
Db      374 TGTGACTCCCAATCTCTTAACAAC----- 397
Qy      208 GluPheSerThrHisGlnGlnValLeuGluGlnGlnMetAsnCysGlyMetGlyAsnGlu 227
Db      398 -----AGTACGCTTCATGAGCTTCAAGAGAGTCCACAT----- 433
Qy      228 ArgAsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSer 247
Db      434 -----CTACAGTCACCTTACTTGTCTCC 454
Qy      248 MetSerProGlySerGlnSerSerCysValThrAlaProSerGlyThrAspSerValAla 267
Db      455 ATGGGACCAACCGCTGGTAATATGTTGTAGACAAAGCTTCAACATCCGAGACCCGAG 514
Qy      268 ValAspAlaValSerGlyAsnGlyHisAlaValLeuGlyGlnLeuGlnProValHisArgLys 287
Db      515 GATAACGCTAGCGGTGAGACACTAGCCGTGTGAGACGGCCACGCCA-----AGACGT 568
Qy      288 SerLeuAspThrPheGlyGlnArgThrSerGlnThrArgGlyValThrArgHisArgTrp 307
Db      569 GCATTGACACTTTCGACAAACGACCTCATCTATGCTGTGTCACAAACATCATAGG 628
Qy      308 ThrGlyArgTrpGlyAlaHisLeuTrpAspAsnSerCysLysGlyGlnGlnThrArg 327
Db      629 ACTGCTGATGTAGGCTCATCTATGAGATATAGTTGTAGAGGAGGCCAGCTCAGG 688
Qy      328 LysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGluGluLysAlaAlaArgAlaTyr 347
Db      689 AAAGAGAAACAAAGTTTACTTGGGTGATGATGACAAAGATTAAGAGCAAGATCATATA 748
Qy      348 AspLeuAlaAlaLeuLysTyrTrpGlyProSerThrHisLeuAsnPheSerLeuGluAsn 367
Db      749 GATCTAGCTGACCTTAAGTACTGGGTCCTTCAACTCTACTAATTTCCCATTTTACAAC 808
Qy      368 TyrGlnValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyrValAlaHisLeu 387
Db      809 TACGAGAAAGAGTAGAGGAAATGACACATGACGAGACAGAGGTTCGGCTGCCATT 868
Qy      388 ArgArgLysSerSerGlyPheSerArgGlyValAserLeuTyrArgGlyValThrArgHis 407
Db      869 AGAAGAGAAAGTAGATGATTTTCAGAGGCCCTTCAGATGTCAGAGAGATTACAGGCAT 928
Qy      408 HisGlnHisGlyArgTrpGlnAlaArgLysGlyArgValAlaGlyAsnLysAspLeuTyr 427
Db      929 CACCAACATGAGAGTGGCAAGCAAGATCGCCGAGTCGCCGAGAAACAAAGACCTCTAC 988
Qy      428 LeuGlyThrPheSerThrGlnGluGlnAlaAlaGluAlaTyrAspValAlaAlaLeuLys 447
Db      989 TTGGGAACCTTTAGACTGAGAGAGAGACACAGAGCTTACGATATAGCTGCATATAAG 1048
Qy      448 PheArgGlyAlaAsnAlaValThrAsnPheAspLysSerArgTyrAspValGluArgLys 467
Db      1049 TTTAGAGAGACTTAATGACAGTACCACTTCAGATCAACCGGTACGACGTGAAGGCATT 1108
Qy      468 MetAlaSerSerAsnLeu---LeuAlaGlyLeuLeuAlaArgArgLysLysAsp----- 484
Db      1109 CTAGAGAGTAGACTTTCCTCCTAGAGAGCGGACGCTTAACGCTCAAGAGAGCTCAA 1168
Qy      485 -----AsnAspProArgAsnLysAspLysAsp----- 493
Db      1169 GCTCTTAGAGTCTTCAAGAGAAACGAGAGCGAGATGATAGCCCTTGTTCAAGTTCCAG 1228
Qy      493 ----- 493

```

```

Db      1229 TACGCTGCTGCTCGAGACAGAGCTCTGGCTCCACCTCATCAAGACTTCAGCTCAACCT 1288
Qy      494 TyrAsnLysSerVal-----ValThrSerValAsnAsnGluGluThr 507
Db      1289 TACCTCTTAAGCATTTCAACAACCATTTAGAGCCCTTTTCTATCTCTCAGAACATAGAC--- 1345
Qy      508 ValGlnValGlnAlaGlyAsnAsnAsnAsnGluLeuAsnAspSerGlyTrpLysMetValLeu 527
Db      1346 -----ATCTCTCATTCACAAACAAACATCTCAGATTCCTCC-----TCT 1387
Qy      528 PheAsnHisProSerGln----- 533
Db      1388 TTTAATCCCAATGACTATATCCAGACAACTTCATCTCCACCAACAGACCAACATTAAC 1447
Qy      534 ---GlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLysAlaMetAsnCysGlyAsnTyr 552
Db      1448 TTGCGACAAAGTGTGAGCCAAAC-----TCTCAGCAGCTCTACATGCTATCTCAT 1501
Qy      553 ArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuLeuGlyLysAspSerValGlySer 572
Db      1502 AGCAATCCGCTCTG-----CTTCATGACTGTCTCTACCTCTATGCTGACAC 1552
Qy      573 GlyGlnHisAsnMetLeuAspGluSerSerLysIleGlyThrHisPheSerAsnThrSe 592
Db      1553 AATAATTAACAATGAGAGCTTAGTGGAGCTTACAAACACTGACGACATTTCTTGGAAACCA 1612
Qy      592 rSerLeuValThrSer-----LeuSerSerSerArgGluLysAspProGlyLys 608
Db      1613 GGTATTGTTATTTGGTCCAGCTCGACTGTGTGATGACAGGAGAGTTCCAAACGTTAAA 1672
Qy      608 sArgGlyProSerLeuLeuPheProMetProPheMetGluThrLys-----IleValAs 626
Db      1673 ACAG-----ATTACGATATAGCTTCCAGTATGGAACCGAGGAGGTATATAGTGTGG 1723
Qy      626 nProfile 628
Db      1724 ACCAGTG 1730

```

Search completed: March 13, 2004, 05:51:51  
Job time : 486 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 13, 2004, 00:44:08 / Search time 4685 Seconds  
(without alignments)  
6133.710 Million cell updates/sec

Title: US-10-024-632-2

Sequence: 1 MKXINESNTDGNHNLG.....RPSPAISLSHLPFASWTDT 663

Scoring table: BLOSUM62  
Xgapop 10.0 / Xgapext 0.5  
Ygapop 10.0 / Ygapext 0.5  
Fgapop 6.0 / Fgapext 7.0  
Delop 6.0 / Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=x1h  
-Q=/cgn2\_1/USPFO.spool/US10024632/rnatc.09032004.101207.10765/app.query.fasta.1.839  
-DB=GenBml -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blcsm62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HAPSIZE=500 -MINLEN=200000000  
-USER=US10024632.ccn 1.1.2860 @rnatc.09032004.101207.10765 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESUBERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WASN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

GenBml:\*

1: gb\_ba:\*

2: gb\_hg:\*

3: gb\_in:\*

4: gb\_ov:\*

5: gb\_ov:\*

6: gb\_ov:\*

7: gb\_ov:\*

8: gb\_ov:\*

9: gb\_ov:\*

10: gb\_ov:\*

11: gb\_ov:\*

12: gb\_ov:\*

13: gb\_ov:\*

14: gb\_ov:\*

15: gb\_ov:\*

16: gb\_ov:\*

17: gb\_ov:\*

18: gb\_ov:\*

19: gb\_ov:\*

20: gb\_ov:\*

21: gb\_ov:\*

22: gb\_ov:\*

23: gb\_ov:\*

24: gb\_ov:\*

25: gb\_ov:\*

26: gb\_ov:\*

27: gb\_ov:\*

28: gb\_ov:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3502	100.0	2344	6	AX555216 Sequence
2	1906	54.4	2546	8	AY461432 Nicotiana
3	1715	49.0	2323	6	AX555218 Sequence
4	1408.5	40.2	1926	6	AX555220 Sequence
5	1408.5	40.2	2510	6	AX106306 Oryza sat
6	1349.5	38.5	2010	6	AX555223 Sequence
7	1328	37.9	1699	8	AY117207 Arabidops
8	1328	37.9	1905	8	AY117207 Arabidops
9	1328	37.9	2056	8	AY080706 Arabidops
10	1328	37.9	2148	6	BD274516 Methods f
11	1328	37.9	2148	6	AR316367 Sequence
12	1328	37.9	2148	6	AR427901 Sequence
13	1328	37.9	2148	6	AY40256 Arabidops
14	1328	37.9	1738	6	BD274518 Methods f
15	1307	37.3	1738	6	AR316369 Sequence
16	1307	37.3	1738	6	AR427903 Sequence
17	1307	37.3	1738	6	AT1035709 Arabidops
18	1095.5	31.3	110766	8	AT1035709 Arabidops
19	1095.5	31.3	198493	8	AT1035709 Arabidops
20	977	27.9	152684	2	AX555222 Sequence
21	977	27.9	152684	2	AX555222 Sequence
22	949.5	27.1	1440	8	AC133450 Oryza sat
23	947	27.0	1992	8	AF317905 Brassica
24	947	27.0	2011	6	AX056889 Sequence
25	944	27.0	2014	6	AX056889 Sequence
26	944	27.0	2014	6	AF317904 Brassica
27	933	26.6	123820	8	AC103891 Oryza sat
28	933	26.6	166072	8	AC135157 Oryza sat
29	914	26.1	1745	8	AK101959 Oryza sat
30	890	25.4	1773	8	ZMWHCPI
31	876	25.0	2153	8	AK111891 Oryza sat
32	850.5	24.3	2164	8	AY062179 Oryza sat
33	762	22.8	1660	8	AK109839 Oryza sat
34	713	20.4	133739	8	AP005309 Oryza sat
35	695	19.8	121675	2	AC144645 Medicago
36	672	19.2	1069	8	AY133880 Arabidops
37	672	19.2	1237	8	AY045915 Arabidops
38	668	19.1	1572	8	AK111617 Oryza sat
39	664.5	19.0	1866	8	AK106769 Oryza sat
40	664.5	19.0	1910	8	AK112088 Oryza sat
41	663	18.6	5151	6	AX056892 Sequence
42	653	18.6	5151	6	AF317907 Arabidops
43	647.5	18.5	33563	8	AF317907 Arabidops
44	647	18.5	81677	8	AB025629 Arabidops
45	630	18.0	105807	8	AC006085 Arabidops

RESULT 1

## ALIGNMENTS



AX555216  
 LOCUS AX555216 2344 bp DNA linear PAT 27-NOV-2002  
 DEFINITION Sequence 1 from Patent WO02059332.  
 ACCESSION AX555216  
 VERSION AX555216.1 GI:25898744  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 REFERENCE  
 1 He, S.S. and Dotson, S.B.  
 Nucleic acid molecules associated with plant cell proliferation and  
 growth and uses thereof  
 Patent: WO 02059332-A 1 01-AUG-2002;  
 Monsanto Technology LLC (US)  
 FEATURES  
 source  
 1..2344  
 /organism="Glycine max"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3847"  
 242..2233  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD58225.1"  
 /db\_xref="GI:25898745"  
 /translation="MKRIINSENTDGDGNHMLGFSLSPHMKKATSAATVPTTPYKS  
 PSOSLSNFGMCGYGENGNHSPLETPVAKDGLCTLELKRSGOTQVWPTSPPL  
 EDPLGATMGTHGHEYSHERGLSLDSIYVNSONARAQNPRLSGPRQGMHVCQHP  
 YVSGIACHGLYQAPLEBETKETVSDSSLMPEMTGLKMWVATPEEFTSQVLEQ  
 QNRCGMGRKGVSLGVSGLSLSHSPGQSCVTPASGDSVAADAKRGHA  
 KLQKQPVHRSYDTFGQRTQYRGVTRHWTGRYEALHMDNSCKEGQTRKQRYVL  
 GGDMEEKARAIDLALAKVGPSTHINFSLENYQVLEEMKMSRQEVYAHIRKSS  
 GFSRGASLYRGVTRHHCGRMARIGRVAGKLDYLGTFTQEEAALYVAALKEFG  
 ANAVTFDISRYDVERHMASNLGELARAKKNDNRNDIDYKSVIVSVNNEEV  
 QVAGNNNNENDESKMYLFNHPGQOQANNGSDOKIMNGVNRNAPFMAIOLDIG  
 IDVSGGQHNMLDESSKICTHPSMTSSLYTSSSRASRSEPKRPSLLPFMPMETKI  
 VNPIGTSTSWLPSTPTVQMPSPAISLSHLVTFASWTDY"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,02e-228 Length: 2344  
 Score: 3502.00 Matches: 663  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 US-10-024-632-2 (1-663) x AX555216 (1-2344)

QY 101 LysLeuGluAspPheLeuGlyGlyValAlaThrMetGlyThrHisGlyTyrGlySerHisGlu 120  
 Db 542 AATTGGAGGAGCTTTTCTAGTGTGTGCTCACTATAGGAACTCCAGATATAGAACCCAGG 601  
 QY 121 ArgGlyLeuSerLeuAspSerLleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsn 140  
 Db 602 AGAGGTTTACGGCTTCAGACAGCATCTATATATACCTCCAAACGACAGGCTTACCCAAC 661  
 QY 141 ArgAspLeuLeuSerGlnProPheArgGlnGlnGlyHisMetSerValGlnThrHisPro 160  
 Db 662 AGAGACCTTCTTTCACACACCTTTCAGGACCAACGTCATATAGTGTCCAAACACACCT 721  
 QY 161 TyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGluThrThr 180  
 Db 722 TATTACTCAGGCTTGTGCTTGCATGTTATATACAGCACCGTTGGAGGAGAAACAACA 781  
 QY 181 LysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGluGlyLeuLys 200  
 Db 782 AAGGAAACGACAGCTGTCCGATTTCAGCTTCTTAATGCTCAATATACAGAGGCTTGAAA 841  
 QY 201 AsnTPValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGluGlnGlnMet 220  
 Db 842 AACTGGGTGCTCCAAACAGGAGTTTCACTCACACGACGAGTTTGGAGCAGCAATG 901  
 QY 221 AsnCysGlyMetGlyAsnGluLysGlnGlyValSerLeuGlySerValGlyCysGlyGlu 240  
 Db 902 AATTGGCGATGGGAGATGAGAAATAGTGTTGTTTATGAGATCTGTGGGGTGTGAGAG 961  
 QY 241 LeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaPro 260  
 Db 962 TTAGCTCTTAAGCTTATCTATGATCTGCTGTTCTCAGTCTAGTGTGCTACTCTCTCT 1021  
 QY 261 SerGlyThrAspSerValAlaValAspAlaLysLysArgGlyHisAlaLysLeuGlyGln 280  
 Db 1022 TCTGGAAACGATTCCTGTTGCTGTGATGCAAGAAAGAGAGGACGCTAAACTTGTTG 1081  
 QY 281 LysGlnProValHisArgGlySerLleAspThrPheGlyGlnArgThrSerGlnTyrArg 300  
 Db 1082 AAGCACTCTGTCATAGAAATCTATCAGACATTTGGGCAAGAACCTCCAGATATGGA 1141  
 QY 301 GlyValThrArgHisArgGlyThrGlyValArgTyrGluAlaHisLeuTyrAspAsnSerCys 320  
 Db 1142 GGTGTCAACAGCATATGATGACCTGTGATGATGAGCGCATTTGTGGGCAATATGTGTGC 1201  
 QY 321 LysLysGluGlyGlnThrArgLysGlyValArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340  
 Db 1202 AAGAGAGAGGCGCAACTNGAAGAGACGACAGAGTGTATTTGGGGGTTATGATATGAG 1261  
 QY 341 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHis 360  
 Db 1262 GAGAAAGCTGCAAGACCTATGATCTCGGGGCTTAACTGTGGGACCTTCAACGAT 1321  
 QY 361 IleAsnPheSerLleGluAsnTyrGlnValGlnLeuGluGluMetLysAsnMetSerArg 380  
 Db 1322 ATTAACCTTTCATAGAGATATACCAAGTCAACTTGAGAAATAGAACCTGAGCGGA 1381  
 QY 381 GlnGluTyrValAlaHisIleuLysArgLysSerSerGlyPheSerArgValAlaSerLle 400  
 Db 1382 CAGGATACGTTGACACATTTGAGAAAGAAAGACAGCGGGTTTCTAGAGTGCTTCAATA 1441  
 QY 401 TyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgLleGlyValArg 420  
 Db 1442 TACAGAGGGGTACAAAGGATACCAACATGAGATGCAAGCGAGATAGGCGAGATT 1501  
 QY 421 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGluGluAlaGluAla 440  
 Db 1502 GCTGGACAAAGACCTTACCTTGGAGCGTTGACACCAACAGAGAGAGCAGCAGAACCA 1561  
 QY 441 TyrAspValAlaAlaLleLysPheArgGlyValAlaAsnAlaValThrAsnPheAspLleSer 460  
 Db 1562 TACGATGTAGCGGCATCAAAATTTGCGGCCCAATGTGACACCACTTTGACATTTTCA 1621  
 QY 461 ArgTyrAspValGluArgGlyMetAlaSerSerAsnLeuAlaGlyGluLeuAlaArg 480

[illegible]

Db 961 AATGAAGTGAAGTTGG-----GGGCAATTAAGCTATTAACCAATATAAT 1008  
 Qy 217 Glu-----GlnGlnMetAsnCysGlyMetGlyAsnGluAsnGlyValSerLeu 233  
 Db 1009 GAAATCTGTAAGTATGCTGTGTCAGAGCTGTGTGGAGTACTCTGGTTCGTGTGT 1068  
 Qy 224 GlySerValGlyCysGlyGlnLeuGlnSerLeuSerLeuSerMetSerProGlySerGln 253  
 Db 1069 GGTACAGTTGGTGTAAATGATTACAGCTTTGAGCTTGCTTAAGAAATCCGTGTCAG 1128  
 Qy 254 SerSerCysValThr-----AlaProSerGlyThrAspSerValAlaValAsp 269  
 Db 1129 TCTACTGTGTACTCCAAAGCAAAATTTCTCTATAGCTGAGATGTTGGCATTTGA 1188  
 Qy 270 AlaIysIysAspGlyHisAlaIysLeuGlyGlnIysGlnProValHisAspGlySerIle 289  
 Db 1189 ACAAGAAAGAGGCTTCTGAAAGAGTGTCTCAAAACCAACCTGTTCTAAGAAATCTATT 1248  
 Qy 290 AspThrPheGlyGlnArgThrSerGlnThrArgGlyValThrArgHisArgTrpThrGly 309  
 Db 1249 GATACATTGGTGCAGAGAACTTCTCAGATAGAGGTGTCACCTAGGCATAGGTGACAGCT 1308  
 Qy 310 ArgIlyGlnAlaHisLeuTrpAspAsnSerCysIysIysGlyGlnThrArgIysGly 329  
 Db 1309 AGGATGAGGCGCATTTGTGGGATATAGTTGCAAGAGAGAGGCGACACTAGAAAGGA 1368  
 Qy 330 ArgGlnValIlyLeuGlyGlyIlyTrpAspMetGlnIlyValAlaIysAlaIysAspLeu 349  
 Db 1369 AGCGAAGTTTATCTTGGGGGCTATGATATGAAAGAGAAAGCTGCAAGAGCATATGATT 1428  
 Qy 350 AlaAlaLeuIysIlyTrpGlyProSerThrHisIleAsnPheSerIleGluAsnIlyGln 369  
 Db 1429 GCTGCACATAAGTATTTGGGACCTTCACTCACTTAATCTTCGTTAGAGAACTTACAG 1488  
 Qy 370 ValGlnLeuGlnIlyMetIlyAsnMetSerArgGlnGlnIlyValAlaHisLeuArg 389  
 Db 1489 AAAGAGCTTGAAGATATGAAATGACCCGGCAAGATATGTTGCACTTATAGAGG 1548  
 Qy 390 LysSerSerGlyPheSerArgIlyAlaSerIleIlyArgGlyValThrArgHisIleGln 409  
 Db 1549 AAAATAGTCAATTTCTCAAGAGGCTTCAATGATATGAGAGAGTCAAGAGCATCATCAG 1608  
 Qy 410 HisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnIlyAspLeuIlyLeuGly 429  
 Db 1609 CATGAAAGTGGCAGAGCAAGATTTGGAAGAGTGGAGGAAACAAGATCTTCACTTGGG 1668  
 Qy 430 ThrPheSerThrGlnGlnIlyAlaIysAlaIysAspValAlaIysIleIysPheArg 449  
 Db 1669 ACTTTCAGACACACAAAGAGAGCTGCAAGACATATGATGTTGCTGCATTAATTCAGA 1728  
 Qy 450 GlyAlaAsnAlaValIlyThrAsnPheAspIleSerArgIlyAspValGluArgIleMetAla 469  
 Db 1729 GGTGTAATGTCTGTCACAACTTTGACATATCGCATCGACGCGAGGAAATCATGCT 1788  
 Qy 470 SerSerLeuLeuAlaGlyGlnLeuAlaArgGlyIlyAspAsnAspProArgAsn 489  
 Db 1789 ACTAATACACTTCCGCTGCGAGAACTTACTAGACAAACAAGAAAGAGACCA----- 1842  
 Qy 490 LysAspIleAspIlyAsnIlySerValIlyThrSerValAsnAsnGlnIlyThrValGln 509  
 Db 1843 -----ATTGATATACATCTCAGT-----GATCAACAAGATGAAAGAGCATGTGTTCAA 1890  
 Qy 510 ValGlnAlaGlyAsnAsnAsnAsnGlnAsnAspSerGlnIlyPlyMetValIleuPheAsn 529  
 Db 1891 AATGACACACACACACACATGCGAATATGTCACAGATTTGGAATATGCTCATAT--- 1947  
 Qy 530 HisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnIlyIleMetAsnCys 549  
 Db 1948 -----CAATCCCAAT-----CCTTCTCTA 1968  
 Qy 550 GlyAsnIlyArgAsnSerAlaPheSerIleAlaGlnIlyAspLeuIleGlyIlyAspSer 569  
 Db 1969 GGAATTTATAGAAACCAACTTTTCCATGGCATTAACAAGATCTATATGCTATGATTGG 2028

Qy 570 ValGlySerGlyGlnHisAsnLeu---AspGlnSerSerIleGlyThrHisPhe 588  
 Db 2029 ATGAATTCACAGCAACATTCATCTTATGATGATCAACAAACAGATTTGTAACATTTT 2088  
 Qy 589 SerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGlnAlaSerProGlyIys 608  
 Db 2089 TCAATGCTTCATCTTGTGTTCCAGCCTTGGCAGATTCAAGAGAACCAAGTCTTATATA 2148  
 Qy 609 ArgGlyProSerLeuIleuPheProMetProPheMetGlnIlyIleValAsnProIle 628  
 Db 2149 AGTCAAGCCTCTTGTAGTCTTTCAGAACCA----- 2178  
 Qy 629 GlyThrSerValThrSerTrpLeuProSerProThrValGlnMetArgProSerProAla 648  
 Db 2179 ---ACAAATGTCAAAGCATGATTTCCCTCT-----GCTCAACTAGGCAATTCCA--- 2226  
 Qy 649 IleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp 662  
 Db 2227 GTTTCATTTCTACACTACCACTTCTTGTGCTTGAATGAT 2268  
 RESULT 3  
 AX555218  
 LOCUS AX555218 2323 bp DNA linear PAT 27-NOV-2002  
 DEFINITION Sequence 3 from Patent WO02059332.  
 ACCESSION AX555218  
 VERSION AX555218.1 GI:25898746  
 KEYWORDS  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 REFERENCE 1  
 AUTHORS He, S. S. and Dotson, S. B.  
 TITLES Nucleic acid molecules associated with plant cell proliferation and growth and uses thereof  
 JOURNAL Patent: WO 02059332-A 3 01-AUG-2002;  
 FEATURES  
 source Monsanto Technology LLC (US)  
 location/Qualifiers  
 1..2323  
 /organism="Glycine max"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3847"  
 25..2022  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAD58226.1"  
 /db\_xref="GI:25898747"  
 /db\_xref="RMTBMBL:CAD58226"  
 /translation="MKSMDNDNADLNQNMWLGFSLSPPQMHNTIVSSHSDPSAAEV  
 VPTSFYHNAPLSVGYFGLAEENVGLYSALPIPLKSDSLVLETLRSQAQAMA  
 TSTPLENPLFGEMGTPHNYECALTEMLPSLDSVYIOPSRDPNNNOYQNHVQ  
 HISTNQOQOQELQAYSTGLRNDMLITGSGKOSQSDNNNLHVNMGGSDADAPVGLK  
 SHEVNFQSAHSEKMTLVPHYEVAAGSGSIGVAYGDLQSLISMSPSSQSSVTS  
 SHRASPAVDVANDTKRGPEKQNDKQIVHRKSTIDTGQSTQVRYGTRPRRWGRYE  
 AHUMNSCKEKGQSRKQVYIGGIDMEKARAPDLAALKTWGSGTINPLEYQN  
 ELBEMKNTRQSVVAHLRKSGSFRGASMTGVRHGHQGWQRIRGVAKNQLYL  
 GFSTQDEAAEAYDIAIKPRGANAIVTFDTRVYVERIKMSNLSLSSELAIRNRETD  
 NETQCIDQHNKPSAYEDTOEALIMHQKSCSEBNDQMVLYVQSQCLEQNPPTIESD  
 RNQCFEVALDMMFHQEVESKARTHYNSPESLATSLSSEKSGSPDRTSLPMLSGMP  
 STASKLILATNPNNVNSWDPSPHLRPAITLIPQVPAAMTDA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,55e-107 Length: 2323  
 Score: 1715.00 Matches: 378  
 Percent Similarity: 65.86 Conservative: 83  
 Best Local Similarity: 54.004 Mismatches: 165  
 Query Match: 48.974 Indels: 74  
 DB: 6 Gaps: 20

US-10-024-632-2 (1-663) x AX555216 (1-2323)

QY 1 MetValArgIleAsnGluSerAsnAsnThrAspAlaSerMetGluVal 20  
 DB 25 ATGAGAGTATGAGAAATGATGACATGCTTATATATCAAAATGGTGGT 84  
 QY 21 PheSerLeuSerProHisMet-----LysMetGluAlaThr 32  
 DB 85 TTCTCATCTCTCTCAATATGATATATAGAGTTTCTTCACTCAACACTCTCT 144  
 QY 33 SerAlaAlaThrValProThrThr-PheTyrMetSerProSerGlnSerHisLeuSerAsn 52  
 DB 145 GCTGTGTAAGTGTCTTCTTACCAAGCTTTTACCAACACTGCTCCA-----CTTAGAGC 158  
 QY 53 PheGlyMetCysTyrGlyVal---GlyGluAsnGlyAsnPheHisSerProLeuThrVal 71  
 DB 199 TATGCTTCTACTATGACTGTAAGCTGAAATGTTGGATTGTATTCAGCTTGGCCATC 258  
 QY 72 MetProLeuLysSerAspGlySerLeuGlyLeuGluAlaLeuLysArgSerGlnThr 91  
 DB 259 ATGCCCTCAATCTGATGCTCTCTCTATGATTGAACTTTAAGCAGATCAAGCA 318  
 QY 92 GlnValMetValProThrSerSerProLysLeuGluAspPheLeuGlyValAlaThrMet 111  
 DB 319 CAAGCAATGCTACTACTTCAACACCAAACTGAGAACTTTAGGTGGGAAAGCCATG 378  
 QY 112 GlyThr---HisGlyTyr-----GlySerHisGlyArgGlyLeuSerLeuAspSer 127  
 DB 379 GGGACCCCTCATCACTACGAAATGATAGTCCACAGAAACAATGCTGTGAGCTTAGACAGT 438  
 QY 128 IleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuSerGlnPro 147  
 DB 439 GTTTTATAC---ATCCAACTCTCAAGCCGATGACCCCAATATATCAAACTCAACAAAC 495  
 QY 148 PheArgGln-----GlnGlyHisMetSerValGlnThrHisProTyrTyr 162  
 DB 496 CATGTTCAACACATTAGACCAACCAACCAACCAACCAAGCTTCAAGCATATATAC 555  
 QY 163 SerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGluTyrThrLysGlu 162  
 DB 556 TCTACCTTGAGAAACCATGATATATATTA-----GAAAGGTCAAAACAA 600  
 QY 183 ThrHisValSerAspCysSerSerLeuMetProGln----- 194  
 DB 601 AGCCAACTTCTGACAAACACATCTTATGTTCAAAACATGGGTGTGATGATCCGCT 660  
 QY 195 MetThrGluGlyLeuLysAsnTyrValAlaProThrArgGluPheSer-----Thr 211  
 DB 661 CCGTTCCTGCTCAAGAGTGG-----GAAAGTGAAGAACTTCAAGTACGATGCA 714  
 QY 212 HisGlnGln-----ValLeuGluGlnGlnMetAsnGlyMetGlyAsnGluArg 228  
 DB 715 CATGAGTCAAAAGATGATGTTCTCATGTGAGGAAATGCTGGTGAATCAAGG----- 768  
 QY 229 AsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerMet 248  
 DB 769 -----TCCATGATGATCAATGGCTATGATGCTGCAATCCTTGAGCTTTCATG 819  
 QY 249 SerProGlySerGlnSerSerCysValThr-----AlaProSerGlyThr 263  
 DB 820 AGTCTTAAGCTCACTGATGACGATGTCACAAAGTCTCAACGCTTCACTGCTGCTGCT 879  
 QY 264 AspSerValAlaValAspAlaLysLysArgGlyHisAlaLysLeuGlyGlnLysGlnPro 283  
 DB 880 GATTTGTTGGCATGATTAAGAAAGAGGGGCTGAAAGGTGACCAAGCAAAAT 939  
 QY 284 ValHisArgLysSerIleAspThrPheGlyGlnArgTyrSerGlnTyrArgGlyValThr 303  
 DB 940 GTTCTAAGAGATCATTAATGACACCTTTGACAAAGAACTCCAGTATAGAGAGTAA 999  
 QY 304 ATGHisArgTyrThrGlyValArgTyrGluAlaHisLeuTyrPAspAsnSerCysLysLysGlu 323  
 DB 1000 AGGCATAGGTGACCTGAGAGATGATGAACTCATCTTTGGACAAACAGCTGCAAAAAGAG 1059

QY 324 GlyGlnThrArgLysGlyValArgGluValTyrLeuGlyGlyTyrAspMetGluGluValAla 343  
 DB 1060 GGGCAAGCAGGAAAGAGAAAGCAAGATTAATCTTAGGGGCTTATGATGAAAGAAAGCT 1119  
 QY 344 AlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHisLeuAspPhe 363  
 DB 1120 GCGAGAGCTTATGATCTAGCGGCACTCAAGATTTGGGAGCCCTCCACTCACTAAACTTT 1179  
 QY 364 SerIleGluAsnTyrGlnValGlnLeuGluGlyMetLysAsnMetSerArgGlnGlyTyr 383  
 DB 1180 CTTTGGAAATATTAACAATTAACCTTGAAGAAATGAAGAAATATCATGACAAAGATAT 1239  
 QY 384 ValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGly 403  
 DB 1240 GTTGTCAATTTAGAAAGAAAGAAAGCAGCGGATTTCTCAAGAGGGGCTTCACTGACAGAGA 1299  
 QY 404 ValThrArgHisHisGlnHisGlyValArgTyrGlnAlaArgIleGlyValAlaGluAsn 423  
 DB 1300 GTAAACAAGCACACCAACATGAAAGTGGCAAGCTCAATGATGATGCTGCTGAAAC 1359  
 QY 424 LysAspLeuTyrLeuGlyThrPheSerThrGlnGluAlaAlaGluAlaTyrAspVal 443  
 DB 1360 AAAAGATCTATCTTGAACCTTTAGTACAAAGAGAAAGACGCTGAACCTATGATAT 1419  
 QY 444 AlaAlaIleLysPheArgGlyValAlaAlaValThrAsnPheAspIleSerArgTyrAsp 463  
 DB 1420 GCTGTATTAATTAATCCAGAGAGCGAATGCTGTAAACCACTTGACATCAACAAGATATGAT 1479  
 QY 464 ValGluArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLys 483  
 DB 1480 GTGGAGAAATATATGACCAAGCAAGCAATCTCTTACAGATGACTTGTAGCGCAACCA 1539  
 QY 484 AspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValThrSerVal--- 502  
 DB 1540 GAGACGGAACAATGAATCACTAGTCATGATCAAAATCAAAATAGCCTTCTCATATGAG 1599  
 QY 503 AsnAsnGluGluThrValGlnValAlaGlnAlaGlyAsnAsnAsnGluAsnAspSerGlu 522  
 DB 1600 GACACTCAAGAAAGCTATTTATTAATGACCAAGAAAGCTGTGAGAGAAATATAT---CAG 1656  
 QY 523 TyrLysMetValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySer 542  
 DB 1657 TGGAGATGCTCTCTAC---CAATCCTCTCAAGCAACTGACAGATCAACAACAAT 1713  
 QY 543 AspGlnLysIleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGln 562  
 DB 1714 GAG-----AGTACAGAACTTAACCAAGTCTTGGCAGGTGCTTTGAGC 1755  
 QY 563 AspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSer 582  
 DB 1756 AACATGTT-----CATCAGAGAGTGAAGAAATCAAGT 1788  
 QY 583 LysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArg 602  
 DB 1789 AAGGCAAGACCAAGTGTCAAAATCTTCTTATTTGGCCACAAAGTTTGGAGCTTCAAGA 1848  
 QY 603 GlnAlaSerProGluLysArgGlyProSerLeuLeuPheProMetProMetGluThr 622  
 DB 1849 GAAAGTACCTGATAGACAGAGCTTGCACAAAGCTCTCTGAAATGCTTCAACCTGATCA 1908  
 QY 623 LysIleValAsnProIleGlyThrSerValThrSerTyrLeuProSerProThrValGln 642  
 DB 1909 AAACATTTGCTACTAATCAATTAACGAATTCCTTGGAGCCCTTCAACCCCATTTAGG 1968  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTyrThrAsp 662  
 DB 1969 -----CCAGCACTTACTTTGCTCAAAATGCCAGTTTGGACCTTGGACAGAT 2016

RESULT 4  
 AX555220 1926 bp DNA linear PAT 27-NOV-2002  
 LOCUS  
 DEFINITION Sequence 5 from Patent WO02059332.

ACCESSION AX555220  
 VERSION AX555220.1 GI:25898748  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 AUTHORS He, S.S. and Dobson, S.B.  
 TITLE Nucleic acid molecules associated with plant cell proliferation and  
 growth and uses thereof  
 JOURNAL Patent: WO 0205932-A 5 01-AUG-2002;  
 Monsanto Technology LLC (US)  
 FEATURES  
 source  
 1..1926  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 1..1926  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD58227.1"  
 /db\_xref="GI:25898748"  
 /translation="WASGGSSNNMIGFSLSPHMPAMEVPSSEPTAAHHHHHPA  
 AAAAGMSPPDSTATCNPLPSPPAAMVAAPSPGYVVGAYGDSGTAGVYSHLP  
 VPIFSDSLCIMEGMPSSPKLEDFICGNSGSDPAITYSGGDAEDASAAVCH  
 HQLVYNTQPLTEAMLEDAAPMEDMAAKNPLTYSACTGEMQPSLSMS  
 PGOSSSCVSAAPQHQMAVVAALAAAGDQSNNDGGEQVGRKGVIGKQKQKQ  
 VRRKSIDTFGRTSQYRGVTRHRTGRYEALMDNSCKKQDQTKQVLYGQDIED  
 KARAYDLAALKYMLSTHINPLENYDEIEEMRMTRQYVAHLRRSGSGFRGAS  
 IYRGVTRHQRWQARIGRVANKDYLGTSTOEAAEAYDIALIKFGILNAVTF  
 DITRDVDKIMSSSLPGEARVKYALEAPDHPYIGRELGATEEASATVGTDMR  
 NVLHSSQQAALACTEATRADLOKMGNGAHSLNGITGYPYESAADBDIVPGKISG  
 IYFNSSSLVLTSLNSREGSPRIGLAVLAKHPTAVSLAAMPKMPAPAAAHVM  
 RPSPAIAMLPVPAWNTDA"  
 ORIGIN  
 Alignment Scores:  
 pred. No.: 2 31e-86 Length: 1926  
 Score: 1408.50 Matches: 340  
 Percent Similarity: 55.058 Conservative: 63  
 Best Local Similarity: 46.458 Mismatches: 152  
 Query Match: 40.224 Indels: 177  
 DB: 6 Gaps: 26  
 US-10-024-632-2 (1-663) x AX555220 (1-1926)  
 QY 13 GYAASHAHIAENTPLEUglyPheSerLeuSerProHisMetLeuMetGluAlaThr 32  
 Db 13 GGGGAGGAGGACAACTGGTTAGCTTCTGCTCTCCGACATGCGGACATGAGAGGTG 72  
 QY 33 SerAlaAlaThrValProThrThr----- 40  
 Db 73 CCGTCTCTCTGAGGACATGCTGTCTCATCATCATCATCATCATCATCATCTGCT 132  
 QY 41 -----Phe 41  
 Db 133 GCTGCTGCTGCTGCGGACCATGCTCTCTCTCCGACAGCGACGACGACCTGCAACTTC 192  
 QY 42 TYMetSerProSerGlnSerHisLeu-----SerAspMetGly 54  
 Db 193 CTCTTCTCTCTCTCTGACGACAGATGCGCTCTCTCACTGCTACTACTACTGCGC 252  
 QY 55 MetCysTyrGlyValGlyLysn-----GlyAspMetHisSerProLeuThrValMet 72  
 Db 253 GGGGCTTACGAGACCGGACCAAGCCGCGGCTCTACTCTGACCTCTCTGTCATG 312  
 QY 73 ProLeuLysSerAspGlySerLeuCysIleLeuGluAlaLeuLysArgSerGlnThrGln 92  
 Db 313 CTTATCAAGTCCATGCTCTCTCTCATCATGAAAGC----- 351

QY 93 ValMetValProThrSerSerProLysLeuGluLysPheLeuGlyValAlaThrMetGly 112  
 Db 352 --ATGATGCGGTCTCTCATCGCCAAAGGCTTCATGGAGTGTGGCAATGGCAGT 408  
 QY 113 ThrHisGluTyrGlySerHisGluLysArgLysSerLeuAspSerIleTyrTyrAsnSer 132  
 Db 409 GGCCAT-----GACCCGCGCACCTACTATAGC 435  
 QY 133 GlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPheArgGlnGlnLys 152  
 Db 436 CAGGGCCAAAGACAGAGAT-----GCAAGCAGGCGGCTTACCAAGAC 480  
 QY 153 HisMetSerValGlnThrHisIleProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGln 172  
 Db 481 CACCAAGCTAGTC-----CCCTAACAAAC-----TACCAAG 507  
 QY 173 AlaProLeuGluGluGluThrThrLysGluThrHisValSerAspCysSerSerLeuMet 192  
 Db 508 CCATTGACCGAAGCAGAGATGCTGCAAGAGCGCGCAGCGCCCAATGAGACGCAATG 567  
 QY 193 ProGlnMetThrGluGlyLeuLysAsnTyrValAlaProThrArgGluPheSerThrHis 212  
 Db 568 -----CGCGCGGCGCAAGAACTTC----- 585  
 QY 213 GlnGlnValLeuGluGlnGlnMetCysGlyMetGlyAsnGluLysArgGlyValSer 232  
 Db 586 -----CTCGTACACAGCTACGCGGCTCTCAAGGACACAGAGATGCGC----- 630  
 QY 233 LeuGlySerValGlyCysGlyGlyLeuGluGlnSerLeuSerLeuSerMetSerProGlySer 252  
 Db 631 -----CAGCCGCTCAGGCTTCCATGAGCCCAAGGCTC 663  
 QY 253 Gln---SerSerCysValThr---AlaProSerGlyThrAspSerValAlaValAspAla 270  
 Db 664 CAGTCAGACAGCGGCTCAGTCGAGCTCCCGACAGCANTCAGCAGATGCGGCTGCTGCT 723  
 QY 270 ----- 270  
 Db 724 GCAGCTGCTGCTGCTGATGACGAGGACCAAGTAAATGACGCTGCGACAGCGCT 783  
 QY 271 -----LysLeuArgGlyHisIleAlaLysLeuGlyGlnLysGlnProValHisArgLysSer 288  
 Db 784 GTGGGAAAGAGAGGGGACCCGGAAGGGGCAAAAGCAGCTGTTACCGGAAAGTGC 843  
 QY 289 IleAspThrPheGlyGlnLysArgThrSerGlnTyrArgGlyValThrArgHisArgTyrThr 308  
 Db 844 ATTGACACGTTTGGCGACGAGACATCGCAGTATAGGGGCGCTCACCGACAGGAGTGGACT 903  
 QY 309 GlyArgTyrGluAlaHisIleuTyrAspAsnSerCysLysLeuGluGlyGlnThrArgLys 328  
 Db 904 GGAAGATAGAACCCCACTCTGGGATACAGTTGCAAAAAGATGGAACAGCAAGAGAG 963  
 QY 329 GlyArgGlnValTyrLeuGlyGlyTyrAspMetGluGluLysAlaAlaArgAlaTyrAsp 348  
 Db 964 GGAAGGCAATATATAGTATAGTGTATGACACTGAAAGAACTCGCGGCGCTATGAT 1023  
 QY 349 LeuAlaAlaLeuLysTyrTyrGlyProSerThrHisIleAsnPheSerIleGluLysTyr 368  
 Db 1024 CTGGCTGCTGTAATCTGGGGGCTATCTTACCATTAATATTTCCCGTTAGAAACTCAG 1083  
 QY 369 GlnValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyrValAlaHisLeuArg 388  
 Db 1084 CAGATGACAGATCAGAGAGATGGAAGATGACAGCAAGCAAGATATTTCCGACTTGAGA 1143  
 QY 389 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 408  
 Db 1144 AGAGAGACGAGGAGTCTCTCGCGTCTTCATTTACCGGGAGTACAGAGCATCAGC 1203  
 QY 409 GlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 428  
 Db 1204 CAGCATGAGAGATGACAGCTCGATGCGACAGGCTGCGCAACAGCATTTGTATTC 1263  
 QY 429 GlyThrPheSerThrGlnGluGluAlaAlaGluAlaTyrAspValAlaIleLysPhe 448

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	REFERENCE	AUTHORS
Db	1264	GGGACCTTTGACGACTCAGAGAGAGACGACGAGGACGATACGACTTGTGCTGCATCAAGTTC						
QY	449	ATGGTGLAALAsnAlaValThrAsnPheAspLLeSerArgTyrAspValGluArgLLeMet						
Db	1324	CGTGGGCTGGAACCGCGTGAAGAACTTTGACATCAACAAGGTACACGCTGGACAAGATCATAG						
QY	469	AlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLysLys						
Db	1384	GAGAGCGAGCTCGCTGCTGCTGCTGCGTGAAGCGACGCGTAAAGGTGAAGGCATCGAGCGACGC						
QY	484	---AspAsnAspProArgAsnLysAspLLeAspTyrAsnLysSerValValThrSerVal						
Db	1444	CCGGACCATTCATGGCCAAATAGGCCCGCGACGCTC						
QY	503	AsnAsnGluGluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGlu						
Db	1480	ACCGAGGAAGCGACGCGCTGCTACTGTACAGGAC						
QY	523	TrpLysMetValLeuPheAsnHisProSerGlnGlnGlnAla						
Db	1519	TGGAGATTCGTGCTC						
QY	539	GlyAsnGlySerAspGlnLysLLeMetAsnGlySerLysAsnTyrArgAsnSerAlaPheSer						
Db	1573	GCAACGCGCAGATCTTCAGAAAG						
QY	559	MetAlaLeuGlnAspLeuLLeGlyLLeAspSerValGlySerGlyGlnHisAsnMetLeu						
Db	1618	---GCTCTCCACGCGCATTTGCTGGGTGACGCTCGAGTCCGCGCAGCGATGACAGATCATC						
QY	579	AspGlnSerSerLysLLe---GlyThrHisPheSerAsnThrSerSerLeuValThrSer						
Db	1675	GTCGCCGGAGGGAAGATCATGTCGATCATCACTTCCTGAACTGCTTTCGCTGGTATCATAGC						
QY	598	LeuSerSerSerArgGluAlaSerProGluLysArgGlyProSerLeuLeuPhePheMet						
Db	1735	CTGAGCAACTCGAAGGAGAGGAGGAGCCCTGAGAGGCTTGCCCTCGCCATGCTCTACGCCAAG						
QY	618	ProPheMetGluThrLysLLeValAsnProLLeGlyThrSerValThrSer						
Db	1795	CAT						
QY	635	TrpLeuProSerProThrValGln						
Db	1834	TGGATGCGCGATGCGCGCGCGCGCGACCTACGATGATGAGGCGCGCGATGCGCATTT						
QY	651	LeuSerHisLeuProValPheAlaSerTrpThrAsp						
Db	1891	---GCTCATCTCCCTGTTTTCGACCGCTGACACAGAT						
RESULT 5	AKI06306	2510 bp mRNA linear PLN 24-JUL-2003						
LOCUS	AKI06306	Oryza sativa (japonica cultivar-group) cDNA clone:002-101-C10, fulll						
DEFINITION	AKI06306	Insert sequence.						
ACCESSION	AKI06306.1	GI:32991515						
VERSION	AKI06306.1	GI:32991515						
KEYWORDS	FLI_CDNA; oligo capping.							
SOURCE	Oryza sativa (japonica cultivar-group)							
ORGANISM	Oryza sativa (japonica cultivar-group)							
REFERENCE	1	The Rice Full-Length cDNA Consortium. National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,K., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,M., Suzuki,K., Li,C., Ohnuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Oono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,X., Tsunoda,Y.						

TITLE	JOURNAL	FEATURES
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice	Science 301 (5631), 376-379 (2003)	Location/Qualifiers
22752273	12869764	1..2510
2 (bases 1 to 2510)		
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoke, T., Horii, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imanura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Mutsaers, T., Nakamura, M., Namiki, T., Narikawa, R., Nishikawa, R., Nishikawa, T., Nomura, K., Numata, K., Numata, R., Onoda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Ose, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shisihiki, T., Sugabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahita, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.		
Direct Submission		
Submitted (27-MUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)		
This clone is one of the 28K full-length cDNA clones from japonica rice.		
URL: http://cdna.dna.affrc.go.jp/cDNA/NIAS_Rice_Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Kishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Onoda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, R., Shisihiki, T. and Yamamoto, M.		
PAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikawa, R., Nishikawa, T., Nomura, K., Numata, R., Onoda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Ose, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shisihiki, T., Sugabe, Y., Sugano, S., Sugiyama, A., Matsubara, K., and Murakami, K.		
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoke, T., Horii, F., Iida, J., Iida, Y., Imanura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kawamata, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Mutsaers, T., Nakamura, M., Namiki, T., Narikawa, R., Nishikawa, R., Nishikawa, T., Nomura, K., Numata, K., Numata, R., Onoda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Ose, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahita, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.		
Location/Qualifiers		
1..2510		
/organism="Oryza sativa [japonica cultivar-group]"		

/mol\_type="mRNA"  
/db\_xref="taxon:39947"  
/clone="002-101-C10"

## ALIGNMENT SCORES:

Pred. No.: 3,3e-86 Length: 2510  
Score: 1408.50 Matches: 340  
Percent Similarity: 55.05% Conservative: 63  
Best Local Similarity: 46.45% Mismatches: 152  
Query Match: 40.22% Indels: 177  
DB: 8 Gaps: 26

US-10-024-632-2 (1-663) x AK106306 (1-2510)

QY 13 G|YAsnAaHtAsnTPrLeuGlyPheSerLeuSerProHtIsmElyMetGluAlaThr 32  
DB 194 GGGGGGAGAGCACTGGTTAGGCTTCTGCTCTCCGACATCCGGCCATGGAGGTG 253  
QY 33 SerAlaAlaThrValProThrThr----- 40  
DB 254 CCGTCCCTCCCTGAGCATGCACTGCTGCTCATCATCATCATCATCATCATCATCTGCT 313  
QY 41 -----Phe 41  
DB 314 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373  
QY 42 TyMetSerProSerGlnSerHisLeu-----SerAsnPhgGly 54  
DB 374 CTCTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433  
QY 55 MetCysTyrGlyValGlyLeuAsn-----GlyAsnPhenIsmSerProLeuThrValMet 72  
DB 434 GGGGCGCTACGAGAACGGAGAACGAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 493  
QY 73 ProLeuLysSerAspGlySerLeuGlyValLeuGluAlaLeuLysArgSerGlnThrGln 92  
DB 494 CCAATGAGTCCGATGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532  
QY 93 ValMetValProThrSerSerProLysLeuGluAspPheLeuGlyValAlaThrMetGly 112  
DB 533 ---ATGATGCGCGTCATCGCCAAAGCTCGAGGATCTTGGGGTGTGGGAATGGGAGT 589  
QY 113 ThnHisGluTyrGlySerHisGluArgGlyLeuSerLeuAspSerIleTyrTyrAsnSer 132  
DB 590 GGGCCAT-----GACCCGCGCCACCTACTATATAC 616  
QY 133 GlnAsnAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPheArgGlnGly 152  
DB 617 CAGGGCCAGAGAGAGAGAGAT-----GCAAGAGGGCGGCTTACACACAC 661  
QY 153 HisMetSerValGlnThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGln 172  
DB 662 CACCACTAGTC-----CCCTACCAAC-----TACCAAG 688  
QY 173 AlaProLeuGluGluGluThrThrThrHisValSerAspCysSerSerLeuMet 192  
DB 689 CCAATTGACGAGAGAGATGCTGCAAGAGGCGGAGCGGCCCAATGGAGAGAGCATG 748  
QY 193 ProGlnMetThrGluGlyLeuLysAsnTyrValAlaProThrArgGluPheSerThrHis 212  
DB 749 -----GGGGCGGCCCAAGAACTT----- 766  
QY 213 GlnGlnValLeuGluGlnGlnMetAsnCysGlyMetGlyAsnGluLysArgGlyValSer 232  
DB 767 -----CTGTATCAGCAGCTACGGCGGCTGCTGAGCAACACAGAAATGCG----- 811  
QY 233 LeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySer 252  
DB 812 -----CAGCGGCTGAGCTCTTCATGAGCCCGGGATCC 844  
QY 253 Gln---SerSerCysValThr---AlaProSerGlyThrAspSerValAlaValAspAla 270

DB 845 CAGTCCACAGAGTGTGATGACGAGCTCCGACAGCATGACGAGATGGGGTGTGCT 904  
QY 270 ----- 270  
DB 905 GCAGTGTGCTGCTGCTGATGAGCCAGGAGAACAGATGACGGTGGCAGCAGCT 964  
QY 271 -----LysLysArgGlyLysAlaLysLeuGlyGlnLysGlnProValHisArgLysSer 288  
DB 965 GTGGGAGAGAGAGGGGAGCCGGAGAGGGGGCCAAAGCAGCTGTTCACCGGAGTCC 1024  
QY 289 IleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHisArgTyrThr 308  
DB 1025 ATTTGACAGTTTGGGAGAGAGATCCGATATGAGGGCGGTACACAGGACAGGTGACT 1084  
QY 309 GlyArgTyrGluAlaHisLeuTyrAspAsnSerCysLysLysGluGlyGlnThrArgLys 328  
DB 1085 GGAAGATATGAGAGCCCATCTGGGATATACATGTTGAAAAGATGACAGACAGAGAG 1144  
QY 329 GlyArgGlnValTyrLeuGlyGlyTyrAspMetGluLysAlaAlaArgAlaTyrAsp 348  
DB 1145 GGAAGGAGAGATATCTAGGTGTTATGACATGAAATGAAAGCTGAGAGGCTTATGAT 1204  
QY 349 LeuAlaAlaLeuLysTyrTyrGlyProSerThrHisIleAsnPheSerIleGluAsnTyr 368  
DB 1205 CTGGCTGGCTGGAATATCTGGGGCTATCTACGATATTAATTTCCCGTTAGAAAATAC 1264  
QY 369 GlnValGlnLeuGluGluMetLysAsnMetSerArgGlnGluTyrValAlaHisLeuArg 388  
DB 1265 CAGATGAGATCGAGAGATGGAAGATGACAGAGGAAATATGTTGGCACTTGAGA 1324  
QY 389 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 408  
DB 1325 AGGAGAGACAGAGGGTTCTTCGCGGTCTTCATCTACCGGGAGATACAGAGCATCAC 1384  
QY 409 GlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 428  
DB 1385 CAGCATGGAATATGGCAAGCTCGGATTTGGAGGGTGTGTCAGACAGAGCTGTATCTC 1444  
QY 429 GlyThrPheSerThrGlnGluGluAlaGluAlaGlyAspValAlaAlaIleLysPhe 448  
DB 1445 GGCATTTTCAACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTTC 1504  
QY 449 ArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGluAlaIleMet 468  
DB 1505 CGTGGCTGAAAGCGGGTACCAATTTGACATCAAGAGTACCAAGTACCAAGATCATG 1564  
QY 469 AlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLysLys----- 483  
DB 1565 GAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1624  
QY 484 ---AspAsnAspProAspAsnLysAspIleAspTyrAsnLysSerValAlaThrSerVal 502  
DB 1625 CCGGACCAATGTCDAATGCGCGGAGCTC-----GGTGGC 1660  
QY 503 AsnAsnGluGluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGlu 522  
DB 1661 ACCGAGAGAGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1699  
QY 523 TrpLysMetValIleuPheAsnHisProSerGlnGlnGlnAla-----Asn 538  
DB 1700 TGGAGAAATGTGCTC-----CATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1753  
QY 539 GlyAsnGlySerAspGlnLysIleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSer 558  
DB 1754 GCAACGGCAGATCTTCAAG-----GGCTTACGTGGTACGGGCACTCG 1798  
QY 559 MetAlaLeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeu 578  
DB 1799 ---GCTTCCACGGCATTTGGGTTGAGTCACTGAGTGGAGGAGAGTCAAGATCATGAT 1855  
QY 579 AspGluSerSerLysIle---GlyThrHisPheSerAsnThrSerSerLeuValThrSer 597



Db 1856 GTCCGGGAGGAGATCATGTCATCACTTCTGCACTCTGCTGCTGACTAGC 1915  
 QY LeuSerSerSerArgGluAlaSerProGluArgGlyProSerLeuLeuPheProMet 617  
 Db 1916 CTGACCACTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1975  
 QY 618 ProPheMetGluThrIleValAsnProIleGlyThrSerValThrSer----- 634  
 Db 1976 CAT-----CATCCACCCCGCTGACGCTCGCCGACATGAAACCC 2014  
 QY 635 ThrLeuProSerProThrValGln-----MetArgProSerProAlaIleSer 650  
 Db 2015 TGAATGCCATGCCCCGCGCCGCGACCTCATGATGAGAGGCGCGGAGTGCATT-- 2071  
 QY 651 LeuSerHisLeuProValPheAlaSerTrpThrAsp 662  
 Db 2072 ---GCTCATCTCCCTGTTTTCAGCCTGACAGAT 2104  
 RESULT 6  
 AX555223  
 LOCUS AX555223 2010 bp DNA linear PART 27-NOV-2002  
 DEFINITION Sequence 8 from Patent WO02059332.  
 ACCESSION AX555223  
 VERSION AX555223.1 GI:25898751  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoideae; Oryza.  
 REFERENCE  
 1 He, S.S. and Dotson, S.B.  
 Nucleic acid molecules associated with plant cell proliferation and  
 growth and uses thereof  
 Patent: WO 02059332-A 8 01-AUG-2002;  
 Monsanto Technology LLC (US)  
 JOURNAL  
 Location/Qualifiers  
 1..2010  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 1..2010  
 /note="unnamed protein product"  
 CDS  
 /codon\_start=1  
 /protein\_id="CAD58228.1"  
 /db\_xref="GI:25898752"  
 /translation="MAGNSSSSSGSMATAGVGWLGPSLPHNATYAGAGVDVVG  
 HHHHHVHQQOHHGGLFYNPAAVVASFYGGHDLVVTSAAGSGSYVAGSSMPL  
 KSDGLCTMEALRGDDRGQGVVVASPKLEDFLAGPAMALSLDNGAFYGGHGH  
 GHADGAVGDPHHGGGGLCAVITGAGAGHDAALVHQAAYAAVAGPAMAGGY  
 DLANADVDVCAAGTTPTGHIHPLTSMSSAGSOSCTTCAAAGAPRYAMDAVS  
 KRGGADRAQKQPVHRSITTFGQRTSQRGTGTRHRTGRTYEAHMDNSCKEGQTR  
 KGRVYLGGYDDEEKAAYDLALKTWGPSTHNPLEDYQELBEMKMSQEVYA  
 HLRKSGSFRSASLYRGVTHQHQWAKIGVSNKQLYGTFIASFAARAR  
 HAGTOEEAAEAYDVAIKFRGLNAVNTDITRYDKILSSSTLLGELARFGKGY  
 GGGAAVADAAALVAQGNVAEMKATLALPAARTEOOQHGGHGHGHDHLLPDA  
 FSVLDIVSTVDAGAPRAAPRAPHMSMAATSLNSRSDPDGVDGAGGGGGLATLPAPK  
 AAASKLYSPVPLNTWASPSPAVSVPRAAGVSIHLMFMAWMDA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,48e-82 Length: 2010  
 Score: 1349.50 Matches: 351  
 Percent Similarity: 57.32% Conservative: 68  
 Best Local Similarity: 48.02% Mismatches: 153  
 Query Match: 38.54% Indels: 161  
 DB: 6 Gaps: 35  
 US-10-024-632-2 (1-663) x AX555223 (1-2010)  
 QY 13 G1yAsnAsnHisAsnTrpLeuGlyPheSerLeuSerProHisMet-LysMetGluAla-- 31

Db 52 GAAAGTGNCGCGGCTGCTGAGATTCCTGCTGCGCCGACATGGGAGACTGCGCC 111  
 QY 32 -----ThieerLAlaThrValProThrThrPheTyMetSerProSerGlnse 48  
 Db 112 GCGCGCTGCGAGATGTGGCGACACACACACACACACACACACACACACACAC 170  
 QY 48 His-----LeuSerAsnPheGlyMetCysTy 57  
 Db 171 GCATGAGAGTGGGTGTGTTCAACACCTGCGCGCGTGCCTCTCTTC-----TACTA 224  
 QY 57 RGIValGlyValAsn-----GlyAsnPheHis--Se 67  
 Db 225 CCGCGCGCGGATAGACCGCGCTGACCTTCGCGCGCGCGGAGATGACTATGCGCG 284  
 QY 67 rProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeuGluAlaLeuTy 87  
 Db 285 CCGGTTCTCTCCATGCGCGCTCAAGTCCGACGCGCTGCTGATGATGAGAGCACTCCG 344  
 QY 87 s-----ArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGluAsp 105  
 Db 345 GAGAGCGACCAAGAACAGACAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 404  
 QY 105 eLeuGly---GlyAlaThrMetGlyThrHisGlyTyGlySerHisGlyAlaGlyLeu 124  
 Db 405 CTAAGCGCGCGCGCGCGCATG-----GCGCTGAG 434  
 QY 124 rLeuAsp-----SerLeuTyTyAsnSerGlnAsnHisGluAlaGlnProAsnArg 142  
 Db 435 CTTGACACATCCGCTCTTACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 477  
 QY 142 rLeuLeuSerGlnProPheArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 162  
 Db 478 -----CAGGACACCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527  
 QY 162 rYrSerGlyLeuAlaCysHisGlyLeuTyGln---AlaProLeuGluGluGluThr 181  
 Db 528 CCGCGCGCGCTCTGAGATGCGCTGATCGCTGATCGCGCGCGCGCGCGCGCGCGCG 567  
 QY 181 ysgLuthHisValSerAspCysSerSerLeuMetProGlnMetTrpGlnGlyLeuAsp 201  
 Db 588 GCTG-GTGCACAGACACAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 625  
 QY 201 snTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGluGlnMet 221  
 Db 626 CTTGCGCGCGCG-----ATGCAC----- 642  
 QY 221 snCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCys----- 238  
 Db 643 -----GGCGCGCGCTACAGACATCGCCACCGCGCGCGCGCGCGCGCGCGCG 694  
 QY 239 -----GlyLeuGlnSerLeuSerLeuSerMet---SerProG 251  
 Db 695 GCCCATATATCCACCG 754  
 QY 251 ySerGlnSerSerCysValThr-----AlaProSerGlyThrAspSerValAla 268  
 Db 755 GGTCCCATGTCAGCTGCGTCAAGCGTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 814  
 QY 268 aAlaPala-----LysLysArgGlyHisAla---LysLeuGlyValLysGlnProVal 285  
 Db 815 TGAAGCGCTGAGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 874  
 QY 285 iArgLysSerLeuAspThrPheGlyGlnArgThrSerGlnTyArgGlyValThrArg 305  
 Db 875 ACCGCAATCATGACACGCTCGCGCGAGAGAGTGCAGTCAAGAGGCGTCAACAGCG 934  
 QY 305 iArgTrpThrGlyArgTyGlyAlaHisLeuTrpAspAsnSerCysLysLysGlyGly 325  
 Db 935 ATAGGTGACTGAGAGATATGAGCGACACTCTGAGCAACACCTGCAAGAGAGAGCG 994  
 QY 325 iThrArgLysGlyArgGlnValTyLeuGlyGlyTyArgPheMetGluGluValAla 345  
 Db 995 AGACCAAGAAAGAGCGCAAGTATCTTGTGTGGTATGACATGAGAGAGAGCGTGC 1054



```

QY 345 rgaiaayraspleuvalaleuayrtirpgrlyproserthnislleaspheseri 365
Db 1055 GGGGGTATGATCTTCTGCGCTCAAGTACGGGGCCCTTCCAGCACATCACTTCCCG 1114
QY 365 legiuanyrgrlvalglneugluwetlyasnmetserrargnglnttyvala 385
Db 1115 TGAGAGACTACAGAGAGAGCTGAGAGAGATGAAGAAATGAGAGAGAGAGATATGTGG 1174
QY 385 lahlaleuayrgrlysergrlypheserarglyalaserlletyrfarglyvalr 405
Db 1175 CTCACTAGAAAGAAAGACAGTGGCTTCTGCGCTGCGCTTCACTTACCGTGAATCA 1234
QY 405 hrarghahleaglhlhlegllyrgrlvalaargllleglyargvalalagllyasnl 425
Db 1235 CCAAGCATCATCAGACAGAGAGATGAGAGAGCGGCGGCGGCTCTCGGCGCAACAAAG 1294
QY 425 spleutyrlenglythrphe----- 431
Db 1295 ACCCTTACTTGGGACATTCATCCGCTGCGCTTGGCGCGCGCGCGCGCGCGCATG 1354
QY 432 --serthnglglnglualaglualayraspvalalalalelyspheargly 451
Db 1355 CCGGACACGAGAGAGAGCGCGGAGCGGTACGACGTGCGCGCATCAAGTCCGCGGCG 1414
QY 451 laasnaalavalthrasnpheserlyserarglyraspvalgluargllewerlaser 471
Db 1415 TCACAGCGCGTCAACCACTTCAACATCAGAGAGTACAGACGTGACAGATCTTGAAGACA 1474
QY 471 erasnleuvalagllyleuvalaargaly--lyaspsasnaspproargant 490
Db 1475 GCAACCTCTCTCCCGGGGAGCGTGGCGCGCGCAAGGTAGGTGCGGACCGCGCGCG- 1533
QY 490 yaspslleasprlyrasnlysserlvaltrnservalasnasnglulthryalgn 510
Db 1534 -----GCGGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCTTGG 1567
QY 510 aiglialagllyasnasnleuvalasnaspserrgrlyrlymetvalleu----- 527
Db 1568 TCGACGCCCGG-----AACGTGCGCGAGTGGAAAGTGGCCACCGCGCGCG 1612
QY 528 -----Pheasnhleprosergrlglnglglnglalaasnglyasnglyserapnt 545
Db 1613 CGCTGCGACGCGCGCGGAGAGAGAGCGACAGACACATGCGACGCGCGCGCACCA- 1671
QY 545 yallleuayrncysgllyasnyrargasnserralapheasermetlaleuaglinspleu 565
Db 1672 -----CACCATGACCTCTCCGCGGAGGAGCGCTTCTCGGTG---CTGACAGACATCG 1720
QY 565 leglly--lleaspserrvalgllysergllyhlhlsasnmelueaspolsuserlyst 584
Db 1721 TGTGCAACGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG- 1755
QY 584 legllythrlhrlspheaseranthrserrserleuvalthrserrseraserarglyua 604
Db 1756 --GCGCGGACATGTGATGCGCGCG-----ACGAGCTCGGCAACTCCCGGAGAC 1804
QY 604 laserproglulnyasarglypro-----Serleuueup 615
Db 1805 AGAGCCCTGAC--AGGGGCGTGGCGCGCGCGCGCGCGCGCGCTCTCGCACAGCTGT 1861
QY 615 heyrmetproprometglulthrylalevalasnprollleglythrservalthrsert 635
Db 1862 TCGCCAGCGCGCGCGCGCGCTGAGCTGACAGCGCGGTG-----CGCTGAACACT 1915
QY 635 rpleu--Proserprothvalglmeafargproserprola--lleasrleuser 653
Db 1916 GCGCGCTCGCGCTCGCGCGCGGTGAGCTCGGTGCGCGGAGCGCGCGGTGTCATCGCGC 1975
QY 653 lsleupovalpheaalserttrpasp 662
Db 1976 ACCTGCCAATGTTCGCGCGTGAACCGAC 2004

```

```

RESULT 7
AY117207
LOCUS
DEFINITION
Arabidopsis thaliana putative ovule development protein
accession
AY117207
VERSION
AY117207.1
KEYWORDS
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 1699)
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, B., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1699)
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, B., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (03-JUN-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL CDNA (RAFL CDNA: 'RIKEN
Arabidopsis Full-Length CDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PGEC (SFP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL CDNA: Yamada, K.,
Banh, J., Chan, M.M., Chang, C.H., Chang, B., Dale, J.M., Deng, J.M.,
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,
Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,
Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Yamada, K. (SFP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SFP
/PGEC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to Genbank.
FEATURES
source
location/Qualifiers
1..1699
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U11873"
/ecotype="Columbia"
/notes="This clone is in pUNI 51."
1..1699
/gene="At4g37750"
1..1668
/gene="At4g37750"
/codon_start=1
/evidence=experimental
/product="putative ovule development protein atnegenmenta"

```

```

/protein_id="PAM51282.1"
/db_xref="GI:21436287"
/translation="MKSPCDNDNNNSNTTNTLLGFLSSNNMKMGGRGREAIVSSST
/translation="MKSPCDNDNNNSNTTNTLLGFLSSNNMKMGGRGREAIVSSST
SSAATSSSVLPOLVVDNNTSNGVCGYSGNPGIYSHSVNPLRSDGLCMEALNR
SSSHNNHODSPRYEDPFGTHHNTSHKEMDLSDLPLNTTHENTTNTNCEFFS
POTNRHEERTRANDPDLTRGGSFVGVYGERQSLSTSNRPGSQSCITGSHHOC
NORQNHOSQNHQOISELVEISVGETTTTAAAKRGGEDVYVQOKIVRKSIDT
FGORTSGYRVTNRHMYRGRYEALMDNNSFKRGRQVYVIGYDDEKARAVDL
AALKYMGPSHTNTSALNKYKEIEMKMTMREYVAHLRKSISGSRGASIVRVTFR
HOGRMQARIGRVAGNDLYLGTFGTQOEAEVAVLAIKFTGNATVTFDIYRYDWD
RIMSNTLLSGFLARNNNSIVVRNTEDOTALNAVVGSGNSKESVTEPRLSFPALFA
LPQVNCMPGSSNMGNNSPMTSPNMEKTVALLTLPMQVFPAMADS"
1669..1699
/gene="At4g37750"

```

```

ORIGIN
3 UTR
Alignment Scores:
Pred. No.: 5 71e-81 Length: 1699
Score: 1328.00 Matches: 312
Percent Similarity: 52.84% Conservative: 69
Best Local Similarity: 43.27% Mismatches: 116
Query Match: 37.92% Indels: 224
Gaps: 21

```

```

US-10-024-632-2 (1-663) x AY117207 (1-1699)
QY 1 MetLysArgLeuGlnGluSerAspAsnThrAspAspGlyAsnAsnHisAsnTrpLeuGly 20
DB 1 ATGAAGCTCTTTTGGATATGATGATATATATATATATATATATATATATATATATATAT 60
QY 21 PheSerLeuSerProHis---MetLysMet----- 29
DB 61 TTCTATGCTCTTCAATATATATATATATATATATATATATATATATATATATATATATAT 120
QY 30 -----GluAlaThrSerAlaAlaThrValProThrThrPheTyMet 43
DB 121 TCATCTTCACTCTTCAGCTGCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 44 SerProSerGlnSerHisLeuSerAspPheGlyMetCysTyArgValGlyGluAsnGly 63
DB 181 GGT-----GACACACATGACACATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 231
QY 64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeu 83
DB 222 GGAATCATCTTCTACATGCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 291
QY 84 GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGlu 103
DB 292 GAAAGCTCTCAACAGATCTTCTCACTCGAATCACCATCAAGATTCAATCTCCCAAGAGTGGAG 351
QY 104 AspPheLeuGlyGlyAlaThrMetGlyThrHisGlyTyArgIleSerHisGlyValArgGly 122
DB 332 GATTTCTTTGGG-----ACCATCAACACACACACACACACACACACACACACACACAC 399
QY 123 ---LeuSerLeuAspSerIleTyArgSerGlnAsnAlaGluAlaGlnProAsnArg 141
DB 400 GATCTTAGCTTAGATGATTATTCTTCAACACACACTCTCAT-----GACCCCAACACG 450
QY 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln 157
DB 451 ACTACCAACTTCAACAGATCTCTTACCTTCCCTCAACACCAAGAACCAAT----- 498
QY 158 ThrHisProTyArgSerGlyLeuAlaCysHisGlyLeuTyArgAlaProLeuGluGlu 177
DB 499 -----GAGGAA 504
QY 178 GluThrThrThyGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu 197
DB 505 GAAACT-----AGAAATTACGGGAATAC-----CTAGATTGACACAT 543
QY 198 GlyLeuLysAsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu 217
DB 544 GGA----- 546

```

```

QY 218 GlnGlnMetAsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 237
DB 547 GGGCTTTTAAATGTAAGGGGTATATGGGAAATTTCAA----- 582
QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerCysVal 257
DB 583 -----CACTACTAGCTTATTCATGAGCCCTGGGTGACATCTAGCTGATC 630
QY 258 ThrAlaProSer----- 261
DB 631 ACTGGCTCTCACACACACCAACCAAAACCAAAACCAAAACCAAAACCAAAACCAACAG 690
QY 262 -----GlyThrAspSerValAlaVal 268
DB 691 CAGATCTTGAAAGCTCTTGTGGAGACAAAGCTGGGTGTGAGACGACGACAAATGCCGCT 750
QY 269 AspAlaLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285
DB 751 GCGAAGAAAGAGAGGGGACAAAGAGATGTTGATGTTGGTGCAGAAACAGATTGTTGAT 810
QY 286 ArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyArgGlyValThrArgHis 305
DB 811 AGAAATCTATCGATCTTTGGACACACACTTCTCAATACCGAGGCGTTACAAAGACAT 870
QY 306 ArgTrpThrGlyArgTyArgValAlaHisLeuTrpAspAsnSerCysLysLysGluGln 325
DB 871 AGATGACGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
QY 326 ThrArgGlyGlyArgGlnValTyArgLeuGlyGlyTyArgPheMetGluGlnValAlaArg 345
DB 931 AGTAAAGAAAGAAAGCAAGTTTATCTGGAGGTTGATGATGATGATGATGATGATGATGATGAT 990
QY 346 AlaTyArgPheValAlaLeuTyArgTyArgTyArgTyArgTyArgTyArgTyArgTyArg 365
DB 991 GCAATGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
QY 366 GluAsnTyArgValGlnLeuGluGluMetCysLysAsnMetSerArgGlnGluTyArgAla 385
DB 1051 GAGAAATTACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1110
QY 386 HisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyArgGlyValThr 405
DB 1111 CATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170
QY 406 ArgHisLeuGlnHisGlyValArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 425
DB 1171 AGACATCACACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1230
QY 426 LeuTyArgLeuGlyThrPheSerThrGlnGlnGluAlaAlaGluAlaTyArgPheValAla 445
DB 1231 CTCTACCTTGAACCTTTGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1290
QY 446 IleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyArgPheValGln 465
DB 1291 ATTAACTTCGGGACCAAAATGCTGACTTACTTGAATATACAGAGATGACATGTTGAT 1350
QY 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgLysLysAspAsn 485
DB 1351 CGTATCATGCTGATGACACACTCTTGTGAGAGATTAGCCGGAAGAAC----- 1401
QY 486 AspProArgAsnLysAspIleAspTyArgAsnLysSerValValThrSerValAsnGln 505
DB 1402 -----AACACAGACATTTGTC-----GTACAGAAATACT 1428
QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGlnTrpLysMet 525
DB 1429 GAA----- 1431
QY 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545
DB 1432 -----GACCAAAAC 1440

```

OY		546	IleMetKsnYscgiAenTYrArGAnSseRAlaphSeSerMeTalaLeuGInSpLeulle	565
Dd		1441	GCTGTAAATGCT-----	1452
OY		566	GlyTlEapSeRValGIsySerGylnHisAmMeTLeuaspClusSerSeryllegly	585
Dd		1452	-----	1452
OY		586	ThrlsPheSeRaentHrSerSerLeuValThrSerLeuSeSerSerAArgIuaLaser	605
Dd		1453	-----GTTGTGAAGCGTGTCCTCCACCAAGAAGTCACT	1485
OY		606	ProGUlsARAgGlyProSerLeu-----LeuPheProMetProPrometGlnTrHyS	623
Dd		1486	ACTCCGAGAGACCTTGAGTTTCGCCGCACATTTCGGTGCCTCAGTAATACAAAAG	1545
OY		624	IlleVal---AsnPcIIleglyThrSeRValThSerTriPneuroProSerProctHrnValGln	642
Dd		1546	ATGTCGGATCAATAATGGCGGGAATATGATGCTCTTGACATCAACAACCCTAATGCTGAG	1605
OY		643	MetaGrProSerProAlalaseSerLeuSerHisLeuProValPhealaseTriPhrAsp	662
Dd		1606	CTTAGA---ACCGTCGCTCTTACTTTCGCTCAGATGCGGTTTTGCGTGTGGGCGCAT	1662
OY		663	Thr 663	
Dd		1663	TCT 1665	
RESULT 8				
ATU41339		1905 bp	mRNA	linear
LOCUS				PLN 23-OCT-1996
DEFINITION	Arabidopsis thaliana ANT (AINTEGUMENTA) mRNA, complete cds.			
ACCESSION	U41339			
VERSION	U41339.1	GI:1244707		
KEYWORDS				
SOURCE	Arabidopsis thaliana (chale cress)			
ORGANISM	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Eurakypota; Magnoliophyta; eudicotyledons; core eudicots;			
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.			
REFERENCE	1 (bases 1 to 1905)			
AUTHORS	Elliot,R.C., Bennett,A.S., Hutner,E., Oakes,M.P., Tucker,W.Q.,			
	Genesys,D., Perez,P. and Smyth,D.R.			
	AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic			
	roles in ovule development and floral organ growth			
	Plant Cell 8 (2), 155-168 (1996)			
TITLE				
JOURNAL	96351414			
MEDLINE	8742707			
PUBMED				
REFERENCE	2 (bases 1 to 1905)			
AUTHORS	Smyth,D.R.			
	Direct Submission			
	Submitted (27-NOV-1995) David Smyth, Genetics and Dev. Biology,			
	Monash University, Wellington Road, Clayton, VIC 3168, Australia			
TITLE	Location/Qualifiers			
JOURNAL	1..1905			
FEATURES				
source	/organism="Arabidopsis thaliana"			
	/mol_type="mRNA"			
	/strain="Landsberg erecta"			
	/db_xref="taxon:3702"			
	/chromosome="4"			
	/map="between spt and cer-9"			
	1..1905			
	/tissue_type="flower; ovule"			
	/gene="AINTEGUMENTA"			
	38..1705			
	/gene="AINTEGUMENTA"			
	/note="similar to APETALA2 protein encoded by GenBank			
	Accession Number U12546"			
	/codon_start=1			
	/product="ANT"			
	/protein_id="AA017364.1"			
	/gb_xref="GI:1244708"			
	/translation="MKSFDNDNNHSNTTNILGPSLSNMKKMGGRGREATYSSST			
gene				
CDS				

	misc_feature	LPOVNRKTRFGSGNMGANNSPWTSNPNELKVALTLDPQMPVPAAMADS 878..1106 /gene="A1NTEGUMENTA" /note="encodes first Ap2 domain"			
	misc_feature	1184..1391 /gene="A1NTEGUMENTA" /note="encodes second Ap2 domain"			
ORIGIN					
Alignment Scores:					
Prod. No.:	6,66e-81	Length:	1905		
Percent Similarity:	1328..00	Matches:	312		
Best Local Similarity:	52.84%	Conservative:	69		
Query Match:	43.27%	Mismatches:	116		
	37.92%	Indels:	224		
	8	Gaps:	21		
US-10-024-632-2 (1-663) x ATU41339 (1-1905)					
QY	1 MetLysArgILaSuGluSerAsnSerAsnThrAspAspArgLysAsnSerHisAsnThrLeuGly	20			
DB	38 ATGAGACTCTTTTGTGATATGATGATATAATGATACAGCAACAGATTAATTGTTAGGG	97			
QY	21 PheSerLeuSerProHis::MetLysMet::	29			
DB	98 TTTCATTGCTTCAAATGATGATAAATGGAGGTGAGAGGAGGTAGAGAGCTATTAC	157			
QY	30 ::::GluAlaHisSerAlaIaIaThrValProThrThrPheTyMet	43			
DB	158 TCATCTTCAACTTCTTCACCTGCACACTCTTCTTCTGTCACCTCAACTGTTGTT	217			
QY	44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGlyGluAsnGly	63			
DB	218 GGT-----GACAACTAGCAACTTGGTGGTTCATGGATCTAACCCAAATGGA	268			
QY	64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeu	83			
DB	269 GGAATCATCTCAACATGCTGTGATGATGCACATCAACATTTGATGTTCTTTGCTTATG	328			
QY	84 GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGlu	1030			
DB	329 GAACCTTCAACAGATCTTCTCACTCGAATCAACCATCAAGATTATCTCCAAAGGTGAG	388			
QY	104 AspPheLeuGlyValaIaThrMetGlyThrHisGlyTyrGlySerHisGluArgGly---	122			
DB	389 GATTTCCTTTGGG-----ACCATTCACAAACAAACACAAAGTCAACAAAGAACCATG	436			
QY	123 ---LeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsnArg	141			
DB	437 GATCTTAGCTTAAGTATTATTATTTACAAACACACTACT-----GACCCAAAGC	487			
QY	142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln	157			
DB	488 ACTCAAAACCTTCAGAGGCTTTAGCTTCCTCTCAACCAACAGAACCAT-----	539			
QY	158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGlu	177			
DB	536 -----GAGGA	541			
QY	178 GluThrThrLysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu	197			
DB	542 GAAACT-----AGAAATTACGGGAATGAC-----CTAGATTGACACAT	580			
QY	198 GlyLeuLysAsnThrValaIaProThrArgGluPheSerThrHisGlnGlnValLeuGlu	217			

```

Db      581 GGA----- 583
Qy      218 GInGInMetAsnCySGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 237
Db      584 GGGCTTTTATGATGAGGGTATATGGGAATTTCAA----- 619
Qy      238 CySGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCyVal 257
Db      620 -----CAGTCACCTGAGCTTATCCATGAGCCCTGGGCGACAAATCTAGCTGACATC 667
Qy      258 ThrAlaProSer----- 261
Db      668 ACTGGCTCTCACCAACCAACAAACCAAAACCAACCAACCAACCAACCAACCAACCAAC 727
Qy      262 -----GlyThrAspSerValAlaVal 268
Db      728 CAGATCTCTGAGGCTTTGTGAGACAAAGCGTTGGGTTGAGACGACGACAAATGGCGGCT 787
Qy      269 AspaAlaIysIysArgGlyHis-----AlaIysLeuGlyGlnIysGlnProValHis 285
Db      788 GCGAAGAAGAAAGAGGAGCAAGAGATGTTGATGTTGGTACAGAAACAGATTGTTCA 847
Qy      286 ArgIysSerIleAspThrPheGlyGlnArgThrSerGlnIysArgGlyValThrArgHis 305
Db      848 AGAAATCTATCGATCTTTTGGACAAAGAACTTCTCAATACCGAGCGCTTACAGACAT 907
Qy      306 ArgTrpThrGlyArgGlyValAlaIysLeuThrAspAsnSerCysIysIysGlnGlyGln 325
Db      908 AGATGAGACTGGTATGATGAAAGCTCATCTATGGACAAATGTTCAAGAAAGAGGCTCAC 967
Qy      326 ThrArgIysGlyArgGlnValIysLeuGlyGlyIysAspMetGlnIysAlaAlaArg 345
Db      968 AGTGAAGAAAGACAAAGATTATCTGGAGCTTATATATGAGAGAAAGCTGCTCGCA 1027
Qy      346 AlaTrpAspLeuAlaAlaIysIysTrpGlyProSerThrHisIleAsnProSerIle 365
Db      1028 GCATATGATCTTCTGCACTCAAGTACTGGGGTCCCTCTCTCAACCAATTTCTCTGCG 1087
Qy      366 GlnIysTrpGlnValGlnLeuGlnGluMetIysAsnMetSerArgGlnIysIysValAla 385
Db      1088 GAGAAATTCAGAAAGATGAAAGATGAAAGATGAAAGATGAAAGATGAAAGATGAAAGAT 1147
Qy      386 HisLeuArgArgIysSerSerGlyPheSerArgIysAlaSerIleIysArgGlyValThr 405
Db      1148 CATTGAGAAAGAAAGAGAGAGCTGTTCTCTGAGGGTCTTCCATCTATAAGAGAGTACA 1207
Qy      406 ArgHisIleGlnIleGlyArgTrpGlnAlaArgIleGlyArgValAlaIysAspVal 425
Db      1208 AGACATCCACGACATGAAAGGTGGACACGAGATGGTGAAGTCCGCTGGAAACAAAGAT 1267
Qy      426 LeuTrpLeuGlyThrPheSerThrGlnGlnIysAlaIysAlaIysAlaIysValAla 445
Db      1268 CTCATCTCTGAACTTTTGAAGCCCAAGAAAGAGAGCTTACGATGTCAGCAGCA 1327
Qy      446 IleIysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgIysValGln 465
Db      1328 ATTAAGTCCGTGGCAAAATGCTGTGACTTAACCTTGTATACACAGAGTACAGATGTTGAT 1387
Qy      466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGluLeuAlaArgArgIysIysAspAsn 485
Db      1388 COTATCATCTCTAGTAACACACTCTGCTGAGAGGTTAGCGCAAGAGAAC----- 1438
Qy      486 AspProArgAsnIysAspIleAspTrpAsnIysSerValValIleThrSerValAsnGln 505
Db      1439 -----AACCAAGCATTTGTC-----GTACAGGAATACT 1465
Qy      506 GlnThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnIysAsnAspSerGlnTrpIysMet 525
Db      1466 GAA----- 1468
Qy      526 ValLeuPheAsnHisProSerGlnGlnGlnIysAlaAsnGlySerAspGlnIys 545
Db      1469 -----GACCAACACC 1477

```

```

Qy      546 IleMetAsnCySGlyAsnTrpArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565
Db      1478 GCTCTAAATGCT----- 1489
Qy      566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerIleGly 585
Db      1489 ----- 1489
Qy      586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGlyAlaSer 605
Db      1490 -----GTTGTGAGAGGTGCTTCCAAAGAGTCACT 1522
Qy      606 ProGlnIysArgIysProSerLeu-----LeuPheProMetProPheMetGlnThrIys 623
Db      1523 ACTCCCGAGAGACTCTTACGTTTCCGCGCATTTTCCGCTTCCCTCAAGTTAATCAAAAG 1582
Qy      624 IleVal-----AsnProIleGlyThrSerValThrSerTrpLeuProSerProThrValGln 642
Db      1583 ATGTCGATCAAAATATGGCGGAAATATGAGTCTTGACATCAAAACCTTAATGCTGAG 1642
Qy      643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp 662
Db      1643 CTTAAG--ACCGTGGCTCTTACTTGGCTCGAATGCGGATTTTGGCTTGGGCGCTGAT 1699
Qy      663 Thr 663
Db      1700 TCT 1702

```

```

RESULT 9
ATU44028
LOCUS      1955 bp mRNA linear PLN 30-JAN-1996
DEFINITION Arabidopsis thaliana transcription factor CKC mRNA, complete cds.
ACCESSION U44028
VERSION    U44028.1 GI:1171428
KEYWORDS   Arabidopsis thaliana (thale cress)
SOURCE     Arabidopsis thaliana
ORGANISM   Arabidopsis thaliana
REFERENCE 1. Vergani, P., Morandini, P. and Soave, C.
AUTHORS    Direct Submission
TITLE      Submitted (29-DEC-1995) Paola Vergani, Dept. of Biology 'L.
JOURNAL    Gorini', University of Milan, Via Celoria 26, Milan, 20133, Italy
FEATURES   location/Qualifiers
            1..1955
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            76..1743
            /function="transcription factor; complements a yeast
            protein kinase C 1 mutant; Method: conceptual translation
            supplied by author."
            /codon_start=1
            /product="CKC"
            /protein_id="AA06281.1"
            /db_xref="GI:1171429"
            /translation="MKSFCDDNDNNHSTNLLGFSLSNNMKKGGGREGAIVSSST
            CSAATSSSVLPOLVVDGNTSNFVCCYGSNPNPGIYSMSVPLSDSLCMEALNR
            SSHSNHODSPKVEDPFGTHHNTSHREAMDLSLSLFYNTHTPEPNTTNOEFPF
            PQRHHEETRYGNDPSLTHGSEFNVCVYGFSQSLMSPGSSCITGSHHQ
            NONCHGNOHOIIEALVETSVGPETTTMAAKKRGQEDVVVVGOKIYRKSIDT
            PCQRSGRGTTRHRTWGRYPAHIMDNEFKRGHSRKRCROYLLGGYDMEEKAAAYDL
            AALKTWGSTHTNFSANIKQETIEKMKNTQETVAHARRKSGATVATNPFDTTRDVO
            HQHGWQARIEGVAGNKDLYLTGFGESAEAYVAATKRGTAATVATNPFDTTRDVO
            RIMSNTLISGELARRNNNSIVRNTEQGTALNAVVEGSKNEVSTPERLISFPAIFA
            LPQVNOIKFGSNMGMSFMTSNPNAELKTVALTLPOWFPVPAWADS"

```

## CDS

```

ORIGIN
Alignment Scores: 6.9e-81 Length: 1955
Pred. No.:

```



Df	1621 ATGTTGGATCAATTATGAGCGGAAATAATGACTGTTCGACATCAAAACCCTTAAGTCTGAG	1680
Oy	643 MetAraProbesProcAlaIleSerLeuSerHisLeuProValPheAlaSerTrpThrAsp	662
Df	1681 CTTAAG--AACGGTGCTCTTACTTGTCTCAGAATGCCGGTTTTCGCTTGCGCTGAT	1737
Oy	663 Thr 663	
Df	1738 TCT 1740	
RESULT 10		
AY080706	2056 bp mRNA linear PLN 18-SEP-2002	
DEFINITION	Arabidopsis thaliana putative ovule development protein	
LOCUS	At4g37750) mRNA, complete cds.	
ACCESSION	AY080706	
VERSION	AY080706.1 GI:19310586	
KEYWORDS	FRT CDNA.	
SOURCE	Arabidopsis thaliana ('haie crese)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsie. 1 (bases 1 to 2056) 2 (bases 1 to 2056)	
AUTHORS	Yanada,K., Liu,S.X., Sekano,H., Pham,P.K., Bann,T., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carinci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamtaya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.U., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.	
TITLE	Arabidopsis Full Length cDNA Clones	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2056)	
AUTHORS	Yanada,K., Bann,T., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carinci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamtaya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.U., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.	
TITLE	Direct Submission	
JOURNAL	Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA	
COMMENT	RKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamtaya,A., Sakurai,T., Carinci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.	
FEATURES	The Salx, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yanada,K., Bann,T., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.U., Shim,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.	
SOURCE	Yanada,K. (SSP/PGEC) and Seki,M. (RKEN GSC) contributed equally to this work. Shinozaki,K. (RKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.	
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.		
Location/Qualifiers		
1..2056		
/organism="Arabidopsis thaliana"		
/mol_type="mRNA"		
/db_xref="taxon:3702"		
/chromosome="4"		
/clone="RAFLO6-86-F19 (R11873)"		

gene	/ecotype="Columbia"		/note="this clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SacI insert."	
5' UTR	1..2056		/gene="At4g37750"	
	1..157		/gene="At4g37750"	
misc_difference	2		/gene="At4g37750"	
	158..1825		/note="not present in genomic sequence"	
CDS	/gene="At4g37750"		/evidence=experimental	
	/codon_start=1		/product="putative ovule development protein aintegumenta	
	/protein_id="AAL85024.1"		/translation="MKSPCDMDNNHNTNTNLLGFLSLNNMMKGGRGREALYSSS	
	/db_xref="GI:19310587"		SSAATSSSSVPPOLVNGDNTSNFVCYGSNPNNGIYSHMSVPLBEDSGSLCEALN	
			SSHSNHHQDSSPKVEDPFGTHHNNTSHKEMDLSLSLFLNTTHERNTTNOEFS	
			POTRNHEETRNRYNDPBLTGHGTSFNVGVGEQSLRSGSSQSSCITSSHHQ	
			NONOHQOHQOQIISSEALVETSVGETTTTAAAKRKQGEVVVVGQKQIVRKSID	
			FOQRISQRYGVPRHWRTYRFAHLMDNSPFKEGHSRGRQVLYIGYDMKEKARAYD	
			ALKATKGPSTNTNFSANENYQKEIDEMKMTROEVVAHLRRKSGFSRGASIVRGTVR	
			HOGHGMARIGRVAQNDIYLGTGTQGEBAALVDVAIKFRNTNAVTPNFDLTRVYV	
			RMSNNTTILSGELARRNNNSIVYRNTEDQYALNAVVBGGSNKEVSPPELLSPFAIF	
			LQVQNMKFGSNMGGNNSPMTSPNAELKTVALLTLPPMPVFAMAADS"	
			1826..2056	
ORIGIN	3' UTR		/gene="At4g37750"	
Alignment Scores:	7,39e-81		Length: 2056	
Pred. No.:	1328.00		Matches: 312	
Score:	52.84%		Conservative: 69	
Percent Similarity:	43.27%		Mismatch: 116	
Best Local Similarity:	37.92%		Indels: 224	
Query Match:	8		Gaps: 21	
DB:				
US-10-024-632-2 (1-663) x AY080706 (1-2056)				
QY	1	MeLYsArgIleAsnGluSerAsnAsnThrAphApolysAsnAsnHisAsnTPPLeuGly	20	
DB	158	ATGAAGTCCTTTTGGATGAATAGATGATATATATCATAGCAACAGATAATTGGTTAGGG	217	
QY	21	PhSeSerLeuSerProHis--MetLysMet-----	29	
DB	218	TTCTCATTTGTTCTCAATATGATGAAATGGAGAGTAGAGAGGTAGAGAGCTATTTTAC	277	
QY	30	-----GluAlaThrSerAlaAlaThrValProThrThrPheTYMet	43	
DB	278	TCATCTTCAACTTCTTCAGCTGCAACTCTTCCTTCCTGTCACCTCAACTGTTGGTT	337	
QY	44	SePProSerGlnSerHisLeuSerAspPheGlyMetCysTYRGLyValGlyGluAsnGly	63	
DB	338	GGT-----GACACCACTAGCAACTTTGGTGTTCATGAGATGATTAACCAATGGA	388	
QY	64	AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuGlySileu	83	
DB	389	GGATCTATTATTCACATGCTGTGTATGTCACCTAGATCTATGATGTTCTTTGCTTAATG	448	
QY	84	GluAlaLeuLysSerGlnThrGlnValMetValProHisSerSerProLysLeuGln	103	
DB	449	GAAGCTCTCAACAGATCTTTCATCGATCAATCAATCAAGATTATCTCCAAAGGTGAG	508	
QY	104	AspPheLeuGlyGlyAlaThrMetGlyThrHisGlnTYRGLySerHisGlnATGAGly--	122	
DB	509	GATTTTCTTTGGG-----ACCATCAACAACAACAAGTACCAAGAAGCATG	556	
QY	123	---LeuSerLeuLysSerSerIleTYRThrAsnSerGlnHisAlaGluAlaGlnProAsnArg	141	
DB	557	GATCTTAGCTTGATGATGATTATCTTACCAACCACTACT-----GAGCCCAACAG	607	

QY 142 AspleutenserGlnPrope-----ArgGlnGlnGlnHisMetSerValGln 157  
 DB 608 ACTCAAAACCTTCAAGAGTTCTTACGCTCCCTCAAAACCAAAACCT----- 655  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlu 177  
 DB 656 -----GAGGAA 661  
 QY 178 GluThrThrIysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlu 197  
 DB 662 GAAACT-----AGAAATTACGGAAATGAC-----CTTACTTGACACAT 700  
 QY 198 GlyLeuIysAsnTyrValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlu 217  
 DB 701 GGA----- 703  
 QY 218 GlnGlnMetAsnGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGly 237  
 DB 704 GGGCTCTTAAATGATAGGGGTATATGGGAATTCAA----- 739  
 QY 238 CysGlyGluLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysVal 257  
 DB 740 -----CAGTACCTGAGCTTATCCATGAGCCCTGGGTCAACATCTAGCTGATC 787  
 QY 258 ThrAlaProSer----- 261  
 DB 788 ACTGGCTCTCACCAACCAAAACCAAAACCAACCAACCAACCAACCAAC 847  
 QY 262 -----GlyThrAspSerValAlaVal 268  
 DB 848 CAGATCTCTAAGCTCTTGGAGACAAGCGTTGGTTGAGACAGACAAACAGCGCGCT 907  
 QY 269 AspAlaIysIysArgGlyHis-----AlaIysLeuGlyGlnIysGlnProValHis 285  
 DB 908 GCGAAGAAAGAGGGGACAGAGAGATGTGTAGTTGTGTGCAAAACAGATGTTCAT 967  
 QY 286 ArgIysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
 DB 968 AGAAATCTATCGATCTTCTTGACACAGAACTTCTCAATCCAGAGCGCTTACAGACAT 1027  
 QY 306 ArgTyrThrGlyArgTyrGlyAlaHisIleuTyrAspAsnSerCysIysGlyGln 325  
 DB 1028 AGATGAGCTGCTAATGAACTCATCTATGAGACATAGTTCCAAAGAGGATCGAC 1087  
 QY 326 ThrArgIysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGluValAlaArg 345  
 DB 1088 AGTGAAGAAAGAGACAGATTTATCTGGAGGTTATGATTCAGAGGAAAGCTGCTCGA 1147  
 QY 346 AlaTyrAspLeuAlaAlaIleuIysTyrTyrProSerThrHisIleAsnPheSerIle 365  
 DB 1148 GCATATGATCTTGGCAGCTCAAGTATCGGGGTCCCTCTCACTCACCAATTTCTCGCG 1207  
 QY 366 GluMetTyrGlnValGlnLeuGlnMetIysAsnMetSerArgGlnGluTyrValAla 385  
 DB 1208 GAGAAATTCAGAAAGAGATTGAAGACCTGAAGAACTGATACAGAAATATGTTGCA 1267  
 QY 386 HisIleuArgArgIysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 405  
 DB 1268 CATTGAGAGAGAGAGAGAGAGGTTTCTAGGGGTGTTCTCATCTATGAGAGGATCACA 1327  
 QY 406 ArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnIysAsp 425  
 DB 1328 AGACATCCACAGCAGAGAGGTGGACAGCGATGTTAGAGTGGCTGGAAACAAAGAT 1387  
 QY 426 LeuTyrLeuGlyTyrPheSerThrGlnGlnGluAlaAlaGluAlaTyrAspValAlaAla 445  
 DB 1388 CTCTACCTTGGACCTTTTGAACCCAAAGAGAGAGAGCTTCAGATGTCAGCGCA 1447  
 QY 446 IleIysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGlu 465  
 DB 1448 ATTAAGTTCGGGCAAAATGCTGTGCTAATCTTGTATATACAGAGGTAGAGATGTTGT 1507  
 QY 466 ArgIleMetAlaSerSerAsnIleuAlaGlyIleuAlaArgArgIysIysAspAsn 485

DB 1508 CGTATCATGCTAGTAACACACTCTTGCTGAGAGGTATGAGCGCAAGGAAAC----- 1558  
 QY 466 AspProArgAsnIysAspIleAspTyrAsnIysSerValValThrSerValAlaAsnGlu 505  
 DB 1559 -----AACACACGACTTGTTC-----GTACGAGAAATACT 1585  
 QY 506 GluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGluTyrIysMet 525  
 DB 1566 GAA----- 1588  
 QY 526 ValIleuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnIys 545  
 DB 1589 -----GACCAAAAC 1597  
 QY 546 IleMetAsnGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 DB 1598 GCTCTAATGCT----- 1609  
 QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerIleGly 585  
 DB 1609 ----- 1609  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgIleAlaSer 605  
 DB 1610 -----CTTGTGAGAGGTGTTCCAAACAAAGACTGACT 1642  
 QY 606 ProGluIysArgGlyProSerLeu-----LeuPheProMetProProMetGluThrIys 623  
 DB 1643 ACTCCCGAGAGACTTGTAGTTTCCGCGAGATTTCCGCTGCTCAAGTATATCAAAAG 1702  
 QY 624 IleVal-----AsnProIleGlyThrSerValThrSerIleProLeuProSerProThrValGln 642  
 DB 1703 ATGTTGGATCAATATATGCGGAAATATGAGTCTTGGATCAACAACTTAAGCTGAG 1762  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisIleuProValPheAlaSerTyrThrAsp 662  
 DB 1763 CTTAAG---ACCGTGTCTTACTTGTCCCTCAGATGCCGTTTGGCTGCTGGGCTGAT 1819  
 QY 663 Thr 663  
 DB 1820 TCT 1822  
 RESULT 11  
 BD274516 2148 bp DNA linear PAT 17-JUN-2003  
 LOCUS BD274516  
 DEFINITION Methode for modifying weight of organ, controlling fertility and enhancing asexual reproduction in plant.  
 ACCESSION BD274516 GI:33084284  
 VERSION BD274516.1  
 KEYWORDS JP 2002534078-A/1.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons, core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 2148)  
 Fischer, R.L. and Mizukami, Y.  
 Methode for modifying weight of organ, controlling fertility and enhancing asexual reproduction in plant  
 Patent: JP 2002534078-A 1 15-OCT-2002;  
 THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 Arabidopsis thaliana (thale cress)  
 OS Arabidopsis thaliana  
 PN JP 2002534078-A/1  
 PD 15-OCT-2002  
 PF 07-JAN-2000 JP 2000592392  
 PR 08-JAN-1999 US 09/227421  
 PI ROBERT L FISCHER, YUKIKO MIZUKAMI  
 PC A01H1/00, C12N5/10, C12N15/09, C12N5/00, C12N15/00 CC  
 AINTEGUMENTA (ANT)  
 CC encodes first AP2 domain  
 CC encodes linker region  
 CC encodes second AP2 domain

```

FEATURES
source
1. 2148
/organism="Arabidopsis thaliana"
/mol_type="Genomic DNA"
/db_xref="taxon:3702"

```

Alignment Scores:	
Pred. No.:	7,846-81
Score:	1388.00
Percent Similarity:	52.84%
Best Local Similarity:	43.27%
Query Match:	37.92%
DB:	6
Gaps:	21
Length:	214
Matches:	312
Conservative:	69
Mismatches:	116
Indels:	224
Gaps:	21

```

US-10-024-632-2 (1-663) x BD274516 (1-2148)
Oy 1 MetLysArgLLeaNGlUSeRAspAenPThrAspArgLYuAspAsnHisAsnTPLeuGLy 20
Db 269 ATGAAGCTCTTTTGTTGGTAAATGATGAATAATACATAGCACAACAGCATTAATTTGTTAGG 322
Oy 21 PheSerLeuSerProHis--MetLysMet 29
Db 329 TTCTCATTTGCTTCCAAATATGATGAATAATGGAGGAGAGAGAGAGATAGAGAAGCATTTTAC 388
Oy 30 -----GluAlaThrSerAlaAlaThrValProThrThrPhenylMet 43
Db 389 TCATCTTCAACTTCTTCAGCTGCACACTTCTTCTTCTGTCCACCTCACTGTGTGT 448
Oy 44 SerProSerGlnSerHisLeuSerAenPheGlyMetCysLYrGLYValGLYLuAsnGLY 63
Db 449 GGT-----GACAACACTGACACTTTGGTGGTTGTGTATGATGATCAACCAATGGA 492
Oy 64 AsnPheHisSerProLeuThrValMetPProLeuLysSerAspLYSerLeuCYelleu 83
Db 500 GGAATCAATCTTCAACATGCTGTGATGACCACTCAAGTCGATCGATGCTCTTTGGCTTATG 552
Oy 84 GluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerProLysLeuGln 103
Db 560 GAACCTTCAACAGATCTTCTCACTGATACACATCAAGATTATCTTCCAAAGGTGGAG 619
Oy 104 AspPheLeuGLYglYalatrMetGLYThrHisGLYrYrGLYSerHisGLuArgGLY--- 122
Db 620 GATTTCTTTGGG-----ACCATCAACAACAACAAGTCACAAGAAAGCCATG 667
Oy 123 ---LeuSerLeuAspSerILErYrYrAsnSerGlnAsnAlaGLuAlaGlnProAsnArg 141
Db 668 GATTTAGCTTAAATGATTATTATCTCAACACCACTCAAT-----GAACCCAAACG 718
Oy 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGLYHisMetSerValGln 157
Db 719 ACTACAAACTTTCAGAGGTCCTTTAGCTTCCTCAACCAACCAACCAAT----- 766
Oy 158 ThrHisProTYrTYrSerGLYLeuAlaCysHisGLYLeuTYrGlnAlaProLeuGLYgln 177
Db 767 -----GAGGAA 772
Oy 178 GluThrThrLYrseGLUThrHisValSerAspCysSerSerLeuMetProGlnMetThrGln 197
Db 773 GAAACT-----AGAAATTACGGGAATGAC-----CCATAGTTTACACAT 811
Oy 198 GLYLeuLysAsnTPrValAlaProThrArgGlnPheSerThrHisGLInGlnValLeuGln 211
Db 812 GGA----- 814
Oy 218 GlnGlnMetAsnCysGLYMetGLYAsnGlnLysArgAsnGLYValSerLeuGLYSerValGLY 233
Db 815 GGGATCTTTTAATGATGGGTATATGGGAATTTCAA----- 850

```

QY	238	CyegIgiuLeuNgInserLeuSerLeuSerMetSerProGlySerGInSerSerYVal	257
Db	851	-----CAGTCACTAGACTTATCCATGAGCCTGGGTCACAATCTTAGCTGCATC	898
QY	258	ThrlaProSer	261
Db	899	ACTGGCTCTCACCAACCAACAAACCAAAACCAACCAACCAACCAACCAACAG	958
Db	959	CAGATCTCTGAAGCTCTTGTGGAGACAAACGCTGGCTTTGAGACAGACAGATGGCGCT	1011
QY	269	AspaIaIySLySArGlyHIs-----AlaLyLeuGlyGInLySGLnProValHIs	285
Db	1019	GCGAAGAAAGAAAGAGGGGCAAGAGAGATGTTGATGTTGGTCAGAAACAGATGTTTCAT	1076
QY	286	ArgLySerSerIleAerThrPhesGlyGInAArgHrSerGInTyTArgGlyValThrArgHIs	305
Db	1079	AGAAATATCATCGATACCTTTGGACAAAGAACTTCTAATCCGAGGGCGTTTCAACAGACAT	1133
QY	306	ArgTprHrGlyArgTyGlyAlaHIsLeuTrpAspAsnSerCyLySLeuGInLyGIn	325
Db	1139	AGATGAGCTGGTAGATAGTAAGCTCATCTATAGGCAACATAGTTTCAAGAAGAGAGCTAC	1199
QY	326	ThrArgLySGLyAArgGInValTyLeuGlyGlyTyTAspMetGInGInLySAlaIaArg	345
Db	1199	AGTGAAGAAAGAAAGCAAGATTATCTGGAGAGTTATGATATGAGAAAGAAACCTCTCGA	1251
QY	346	AlaTyTAspLeuAlaIaIaLeuLySTyTrrGlyProSerThrHIsIleAsnPheserIle	365
Db	1259	GCAATATGATCTTGGCGACCTCAGATACGGGGTCCCTCTACTACACCAATTTCTCTGGG	1311
QY	366	GIuSentryGInValGInLeuGInGluMetLySAsnMetSerArgGInGInTyTValAla	385
Db	1319	GAGAAATTTCAGAAAGAGATGAAAGACTGAAAGAACATGACTAGCAAGAAATATGTTGCA	1376
QY	386	HIsLeuArgAGLySerSerGlyPheserArgGlyAlaSerIleTyTArgGlyValThr	405
Db	1379	CATTGAGAAAGAAAGAGAGTGGTTTCTCTAGGGGTGCTTCATCTATAGAGAGTCA	1433
QY	406	ArgHIsHIsGInHIsGlyAArgTyTrrGInAlaArgIleGlyAArgValAlaGlyAsnLySAsp	425
Db	1439	AGACATCACCGACATGMAAGCTGGACACACGAGATTGGTAGAGTGGCTGGAACCAAGAT	1498
QY	426	LeuTyTLeuGlyTyTrrPheserThrGInGInGInAlaIaGlyAlaTyTAspValAlaIa	445
Db	1499	CTTACCTCTTGAACCTTTTGGAACCCGMAAGAGAGCTGAGAGCTTACGATTAGCAGCA	1551
QY	446	IleLySPhearGlyAlaAsnAlaValThrAsnPheserIleSerArgTyTrrAspValGIn	465
Db	1559	ATTAAAGTTCCCTGGCACAAATGCTGTGCTAACTTTGATTATCACAGAGTACCATTTGGAT	1611
QY	466	ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGlyLeuAlaArgArgLySAspAsn	485
Db	1619	CGATATCATGCTTAGTAACACACTCTTGTCTGAGAGTTAAGCGCGAAGGAC-----	1661
QY	486	AspProArgAsnLySAspIleAspTyTrrAsnLySAspValAlaThrSerValAsnAsnGIn	505
Db	1670	-----AACCAACGCAATTGTC-----GTCCAGAAATACT	1691
QY	506	GIuHrValGInValGInAlaGlyAsnAsnAsnAsnGInuAsnAspSerGInTyTrrLySmet	525
Db	1697	GAA-----	1697
QY	526	ValLeuPhesAsnHIsProSerGInGInGInAlaAsnGlyAsnGlySerAspGInLyS	545
Db	1700	-----GACCAAAAC	1708
QY	546	IleMetAsnCySGLyAsnTyTrrArgAsnSerAlaPheserMetAlaLeuGInAspLeuIle	565
Db	1709	GCTCTAAATGCT-----	1721







QY 262 -----GlyThrAspSerValAlaVal 268  
 Db 959 CAGATCTCTGAAGCTCTTGTGGAGACAGCGCTTGAGCGCGCAATGGCGGCT 1018  
 QY 269 AspAlaLysLysArgGlyHis-----AlaLysLeuGlyGlnLysGlnProValHis 285  
 Db 1019 GCGAAGAGAGAGAGAGAGAGAGATGTGTGATGTTGCTGGTCAGAAAACAGATTGTCAT 1078  
 QY 286 ArgLysSerLysLeuArgThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHis 305  
 Db 1079 AGAAATCTATCGATCTTGTGGACAGAACTTCTCAATCCGAGCGCTTCAAGACAT 1138  
 QY 306 ArgTyrPheGlyArgTyrGlnAlaHisLeuTyrAspAsnSerCysLysLysGlnGln 325  
 Db 1139 AGATGAGCTGGTATGATGAGCTCATCTATGGACATATGTTTCAAGAGAGAGCTCAC 1198  
 QY 326 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlnLysAlaAlaArg 345  
 Db 1199 AGTGAAGAGAGAGAGAGAGAGATTTATCTGGGAGGTTAGATAGAGAGAGAGAGCTCTCGA 1258  
 QY 346 AlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyProSerThrHisLeuAsnPheSerIle 365  
 Db 1259 GCATATGATCTTGGCTGCACTCAAGTACTGGGGTCCCTTACATCCACCAATTTCTCTGG 1318  
 QY 366 GlnAsnTyrGlnValGlnLeuGlnLysMetLysAsnMetSerArgGlnGlnTyrValAla 385  
 Db 1319 GAGATATATCGAAGAGAGATGAGACATGATGACATCTAGACAGAGATATATCTCA 1378  
 QY 386 HisLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 405  
 Db 1379 CATTTGAGAGAGAGAGAGAGAGCTGGTTCTCTAGGGGCTCTTCCATCTATGAGAGAGCTACA 1438  
 QY 406 ArgHisHisGlnHisGlyArgTyrPheGlnAlaArgIleGlyArgValAlaGlyLysnLysAsp 425  
 Db 1439 AGACATACACAGAT 1498  
 QY 426 LeuTyrLeuGlyThrPheSerThrGlnGlnGlnAlaAlaGlnAlaTyrAspValAlaAla 445  
 Db 1499 CTACCTGAT 1558  
 QY 446 IleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValAla 465  
 Db 1559 ATTAGATCCGAT 1618  
 QY 466 ArgIleMetAlaSerSerAsnLeuLeuAlaGlyGlnLeuAlaArgArgLysLysAspAsn 485  
 Db 1619 CGATCATGCTGTAGTACACACATCTTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1669  
 QY 486 AspProArgAsnLysAspIleLeuTyrAsnLysSerValValThrSerValAsnAsnGln 505  
 Db 1670 -----MACACAGCATTTGTC-----GTCCAGAAATCT 1696  
 QY 506 GlnThrValGlnValGlnAlaGlyAsnAsnAsnAsnGlnLysAspSerGlnTyrLysMet 525  
 Db 1697 GAA----- 1699  
 QY 526 ValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
 Db 1700 -----GACCAAAACC 1708  
 QY 546 IleMetAsnCysGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 Db 1709 GCTCTAAATGCT----- 1720  
 QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlnSerSerLysIleGly 585  
 Db 1720 ----- 1720  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArgGlnAlaSer 605  
 Db 1721 -----GTTGTGGAAGAGTGTGTTCCAAACAAAGAGAGTCTAGT 1753

QY 606 ProGlnLysArgGlyProSerLeu-----LeuPheProMetProProMetGlnThrLys 623  
 Db 1754 ACTCCCGAGAGACTCTTGAGTTTCCGGCGAGTTTTCGGGCTTCCCAAGATTAATCAAAAG 1813  
 QY 624 IleVal---AsnProIleGlyThrSerValThrSerTyrLeuProSerProThrValGln 642  
 Db 1814 ATGTCGATCAAAATATGAGCGGAGAAATATGATGATCTTGCATCAAAACCTTAAGCTGAG 1873  
 QY 643 MetArgProSerProAlaLeuSerLeuSerHisLeuProValPheAlaSerTyrThrAsp 662  
 Db 1874 CTTAAG---ACCGTCTCTTACTTGTGCTCAAGATCCGGTTTGTGCTTGGGCTGAT 1930  
 QY 663 Thr 663  
 Db 1931 TCT 1933  
 RESULT 14  
 ATU40256 2148 bp mRNA linear PLN 29-FEB-1996  
 LOCUS Arabidopsis thaliana ovule development protein (AINTEGUMENTA) mRNA,  
 DEFINITION complete cds.  
 ACCESSION U40256.1 GI:1209098  
 VERSION U40256  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 2148)  
 AUTHORS Klucher, K.J., Chow, H., Reiser, L. and Fischer, R.J.  
 TITLE The AINTEGUMENTA Gene of Arabidopsis is Required for Ovule and Female  
 Gametophyte Development and is Related to the Floral Homeotic Gene  
 APTERLAP2  
 JOURNAL Plant Cell 8 (1996) In press  
 REFERENCE 2 (bases 1 to 2148)  
 AUTHORS Klucher, K.  
 TITLE Direct Sublimation  
 JOURNAL Submitted (07-NOV-1995) L. Reiser, Plant Biology, UC Berkeley, 111  
 Koshland Hall, Berkeley, CA 94720, USA  
 FEATURES  
 source  
 1..2148  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="4"  
 1..2148  
 /gene="AINTEGUMENTA"  
 259..1936  
 /gene="AINTEGUMENTA"  
 /note="AP2 domain containing protein"  
 /codon\_start=1  
 /product="AINTEGUMENTA"  
 /protein\_id="AAA91040.1"  
 /db\_xref="GI:1209098"  
 /translation="MKSFDNDNDNNHSTNTNLGFSLSNNMKMGKGGGREGRAIVSSST  
 SAAATSSAVPOLVVDGNTSFGVGSNGGIVSGMSMPLRSGVGLIMALNR  
 SSSNNHDDSPKVEDFRGTHNNHSHKENDSLDLPVNTTREPNTTNPGFFSF  
 PQRNHEERTRYGNDPSLTGSGSFNVGYEFQCSLSLSPSSQSCITGSHHQC  
 NONOHSGNQQLSEALVETSVGEFTTMAAKKQEDVVVVGQQLVHRKSLIT  
 FQORTQIRGVTRHRTGRYEAHLMDNSFKEGHSRRGRQVYLLGYMEEGAARAYDL  
 ALKTYWGSPTHTNPSAENYQKEIDMKRMTEQEVVAHLRRKSGFSRGSALTYRCAVTH  
 HOGHMOARIGLVAGNKKDLYLGTGTQEEAAEADVAIAIKRGNTAAVNTPDITRDYD  
 RIMSGWALISGELARRNNNSIVNTEDQTLNNAVVGSGSKVEVSTPERLLISPAIPA  
 LQVNVQKRGSGNMGMGMSPMTSNPAELKTVALLPLQMPVPAWAAS"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.84e-81 Length: 2148  
 Score: 1328.00 Matches: 312  
 Percent Similarity: 52.84% Conservative: 69  
 Best Local Similarity: 43.27% Mismatches: 116  
 Query Match: 37.92% Indels: 224





Db 1192 GCTGGAACAAAGATCTCTACCTTGGAACCTTCGAACTCAAGAAAGCGCGGAGCC 1251  
QY 441 TYRASPVALAIAIAIleIysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSer 460  
Db 1252 TATGATGTAGCAGCTATCAAGTTCCGTGGCACAAACGCGGTGACTTAACCTTGACATAACA 1311  
QY 461 ArgTYRASPVALGluArgGlyMetAlaSerSerAsnLeuValGlyGluLeuAlaArg 480  
Db 1312 AGGTACGATGTTATGCTATAGCTTAAGCTTAACACTCTTGTCTGAGAGATGCTCGA 1371  
QY 481 ArgLYLysAspAsnAspProArgAsnLYAspIleAspTYRAsnLYSerValValThr 500  
Db 1372 AGGAACGCGC-----AACGCACTCGTGTCCGCG 1398  
QY 501 SerValAsnAsnGluGlu-----ThrValGluValGluAlaGlyAsnAsnAsnAsn 517  
Db 1399 AACATTAGCAGCAGAGAGCGCTTTAACCGCTGTGCGAACGCTGTTCCAAATAAGGAA 1458  
QY 518 GluAsnAspSerGluTrpLYMetValLeuPheAsnHisProSerGlnGlnGlnAla 537  
Db 1459 GTG----- 1461  
QY 538 AsnGlyAsnGlySerAspGlnLYIleMetAsnCYsGlyAsnTYRArgAsnSerAlaPhe 557  
Db 1462 -----GGTAGCCCGGAGAGGTTTGAGTTT----- 1488  
QY 558 SerMetAlaLeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMet 577  
Db 1488 ----- 1488  
QY 578 LeuAspGluSerSerLYIleGlyThrHisPheSerAsnThrSerSerLeuValThrSer 597  
Db 1488 ----- 1488  
QY 598 LeuSerSerSerArgGluAlaSerProGluLYArgGlyProSerLeuLeuPheProMet 617  
Db 1489 -----CCGACGATATTGCGCTG 1506  
QY 618 ProPheMetGluThrLYIle-----ValAsnProIleGlyThrSerValThrSerTrp 635  
Db 1507 CCTCAAGTTGGTCCGAGAGATGTCGAGCAAAATGTGTCCGA--AATATGAGTTCTTGG 1563  
QY 636 LeuProSerProThrValGlnMetArgProSerProAlaIleSerLeuSerHisLeuPro 655  
Db 1564 ACTACGAACCTTAATGCTGATCTCAAG--ACCGTTCTTACTGTGCCGCAATGCCG 1620  
QY 656 ValPheAlaSerTrpThrAspThr 663  
Db 1621 GTTTTCGCTGCGTGGCTGATTTCT 1644

Search completed: March 13, 2004, 07:10:13  
Job time : 4737 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 13, 2004, 07:10:41 ; Search time 444 Seconds

(Without alignments)  
5497.521 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502  
Sequence: 1 MKRINSNNNTDDGNHNLG.....RSPRLSLSHLPVFAFWTDT 663

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Rgapop 6.0 , Rgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus\_p2n model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10024632/runat\_09032004\_101210\_10885/app\_query.fasta\_1.839  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=mp -MINMATCH=0.1  
-LOOCL=0 -LOOEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10024632 @cgn1.1.164 @runat\_09032004\_101210\_10885  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	3502	100.0	2344	13	US-10-024-632-1	Sequence 1, Appl1
2	2193.5	62.6	2405	12	US-10-424-599-136491	Sequence 136491, A
3	1744	49.8	1378	12	US-10-425-114-18711	Sequence 18711, A
4	1715	49.0	2323	13	US-10-024-632-3	Sequence 3, Appl1
5	1408.5	40.2	1926	13	US-10-024-632-5	Sequence 5, Appl1
6	1349.5	38.5	2010	13	US-10-024-632-8	Sequence 8, Appl1
7	1328	37.9	2148	14	US-10-059-911-1	Sequence 1, Appl1
8	1322	37.7	1668	14	US-10-059-911-39	Sequence 39, Appl1
9	1319	37.7	1668	14	US-10-059-911-38	Sequence 38, Appl1
10	1307	37.3	1738	14	US-10-059-911-4	Sequence 4, Appl1
11	1233.5	35.2	1359	14	US-10-059-911-30	Sequence 30, Appl1
12	1224	35.0	1422	14	US-10-059-911-126	Sequence 26, Appl1
13	1185	33.8	1231	12	US-10-425-114-11685	Sequence 11685, A
14	1171.5	33.5	1949	12	US-10-424-599-123485	Sequence 123485, A
15	1129.5	32.3	1113	14	US-10-059-911-32	Sequence 32, Appl1
16	1042	29.8	7367	14	US-10-059-911-27	Sequence 27, Appl1
17	977	27.9	7367	13	US-10-024-632-7	Sequence 7, Appl1
18	974.5	27.8	1466	12	US-10-425-114-3280	Sequence 3280, Ap
19	955.5	27.3	834	14	US-10-059-911-28	Sequence 28, Appl
20	953.5	27.2	2128	15	US-10-374-780A-1763	Sequence 1763, Ap
21	951.5	27.2	1818	15	US-10-374-780A-333	Sequence 333, Ap
22	950.5	27.1	831	14	US-10-059-911-42	Sequence 42, Appl
23	947.5	27.1	618	14	US-10-059-911-34	Sequence 34, Appl
24	943.5	27.0	1500	15	US-10-374-780A-399	Sequence 399, Ap
25	943	26.9	2116	12	US-10-424-599-10030	Sequence 10030, A
26	940	26.8	1551	12	US-10-424-599-123483	Sequence 123483, A
27	939	26.8	2088	15	US-10-374-780A-1738	Sequence 1738, Ap
28	930	26.6	2052	15	US-10-374-780A-2495	Sequence 2495, Ap
29	928	26.5	2169	12	US-10-424-599-91337	Sequence 91337, A
30	911.5	26.0	2147	12	US-10-424-599-70555	Sequence 70555, A
31	903	25.8	951	14	US-10-059-911-40	Sequence 40, Appl
32	895	25.6	1479	12	US-10-425-114-14859	Sequence 14859, A
33	891.5	25.5	1941	15	US-10-374-780A-345	Sequence 345, Ap
34	888	25.4	1985	12	US-10-425-114-783	Sequence 783, Ap
35	887.5	25.3	1182	12	US-10-425-114-19324	Sequence 19324, A
36	869.5	24.8	794	12	US-10-425-114-10762	Sequence 20762, A
37	869.5	24.8	806	15	US-10-374-780A-1742	Sequence 1742, Ap
38	861	24.6	525	14	US-10-059-911-36	Sequence 36, Appl
39	762.5	21.6	1074	14	US-10-059-911-1	Sequence 31, Appl
40	695	19.8	1410	12	US-10-424-599-91850	Sequence 91850, A
41	695	19.8	1410	12	US-10-425-114-9718	Sequence 9718, Ap
42	695	19.8	1410	12	US-10-412-699B-1451	Sequence 1451, Ap
43	695	19.8	1410	15	US-10-374-780A-1303	Sequence 1303, Ap
44	688.5	19.7	2042	12	US-10-412-699B-1462	Sequence 1462, Ap
45	688.5	19.7	2042	15	US-10-374-780A-1309	Sequence 1309, Ap

## ALIGNMENTS

RESULT 1  
US-10-024-632-1  
Sequence 1, Application US/10024632  
Publication No. US2002017093A1  
GENERAL INFORMATION:  
APPLICANT: Monsanto Technology LLC  
APPLICANT: He, Steve S.  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
FILE REFERENCE: 38-21(51837)B  
CURRENT APPLICATION NUMBER: US/10/024, 632  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: US 60/257, 896  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 2344  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (242)..(2233)

## OTHER INFORMATION:

US-10-024-632-1

## Alignment Scores:

Pred. No.: 0  
 Score: 3502.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%

Length: 2344  
 Matches: 663  
 Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-10-024-632-2 (1-663) x US-10-024-632-1 (1-2344)

QY 1 MetLysArgIleAsnGluSerAsnThrAspAspGlyAsnAsnHisAsnTrpLeuGly 20  
 Db 242 ATGAAGCCGCAATTAATGAGATGATACACACCGATGATGAAACATCATTAAGTGGGG 301  
 QY 21 PheSerLeuSerProHisMetLysMetGluAlaThrSerAlaAlaThrValProThrThr 40  
 Db 302 TTCTCTCTCTCAACCCACATGAAATGAGGCTACTTCAGCAGCCACTGTTCCAGAAC 361  
 QY 41 PheTyrMetSerProSerGlnSerHisLeuSerAsnProGlyMetCysTyrGlyValGly 60  
 Db 362 TTCTCAAGTCCCTCTTCAATCTCACTGTCACAACTTCGGAATGTGTACGGTGTGCGA 421  
 QY 61 GluAsnGlyAsnProHisSerProLeuThrValMetProLeuLysSerAspGlySerLeu 80  
 Db 422 GAAATAGGTAATCTTCACTTCCACTTACGGTTATGCTTCAAGTCTCAAGTGTGAGGTC 481  
 QY 81 CysIleLeuGluAlaLeuLysArgSerGlnThrGlnValMetValProThrSerSerPro 100  
 Db 482 TGTATCTTGGAACTCTTCAAAAGATCAAAAGCGAAAGTGTGTGCCAACTTCGTCGCG 541  
 QY 101 LysLeuGluAspPheLeuGlyGlyAlaThrMetGlyThrHisGlyTyrGlySerHisGlu 120  
 Db 542 AATATGGAGGACTTCTAGGTGTGCAACTATGAGAACTCAAGATATGAGAGCCACAG 601  
 QY 121 ArgGlyLeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAlaGluAlaGlnProAsn 140  
 Db 602 AAGAGTTTGAAGCTTGACAGCATCTATTATTAATCTCCAAAGCGACAGAGCTCAACCC 661  
 QY 141 AsnAspLeuLeuSerGlnProPheArgGlnGlnGlyHisMetSerValGlnThrHisPro 160  
 Db 662 AAGAGCTTCTTTCACAACTTCAGGCAACAAGTCAATATGATGATGATGATGATGAT 721  
 QY 161 TyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGlnGlnGlnThr 180  
 Db 722 TATTACTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781  
 QY 181 LysGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGlnGlyLeuLys 200  
 Db 782 AAGGAACCGACAGCTGCGGATGACAGCTCCCTAATCCCTCAATGACAGAGGCTTGA 841  
 QY 201 AsnTrpValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGlnGlnGlnMet 220  
 Db 842 AACTGGGTGCTCCAAAGAGGATTTTCACTTCACAGCAAGATTTTGGACAGCAAAATG 901  
 QY 221 AsnCysGlyMetGlyAsnGluArgAsnGlyValSerLeuGlySerValGlyCysGlyGlu 240  
 Db 902 AATTGTGGCATGGGGAATGAGAAATGGTGCTCTTTAGGATCTGTGGGGGTGGAGAG 961  
 QY 241 LeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValAlaThrAlaPro 260  
 Db 962 TTACAGCTCTAAGCTTATATATGAGTCCGCTTCAGCTTCAAGTGTGTCTCACTGCTCT 1021  
 QY 261 SerGlyThrAspSerValAlaValAlaAspAlaLysValArgGlyHisAlaLysLeuGlyGln 280  
 Db 1022 TCTGGAAACAAATCTGTGCTGTGAGTGCAGAAAGAGAGGAGCATCTAAACTGGTCA 1081  
 QY 281 LysGlnProValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyrArg 300  
 Db 1082 AAGCAGCTGTGATAGAAATATCATGACACATTTGGGCAAGAACTTCGCGATATAGA 1141

QY 301 GlyValIleThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCys 320  
 Db 1142 GGTGTCAAGGCAATGATGAGCTGTGATGATAGAGCGCTTTGTGGATTAATGTTGC 1201  
 QY 321 LysLysGlnGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340  
 Db 1202 AAGAGGAAAGGCAACACTAGAGAAAGAGCAAGATGATTTGGGGGCTTATGATATGAG 1261  
 QY 341 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyProSerThrHis 360  
 Db 1262 GAGAAAGCTGCAGAGCCCTATGATCTCGGGCCCTTAAGTCTGGGAGCTTCAAGCAT 1321  
 QY 361 IleAsnProSerIleGluAsnTyrGlnValGlnLeuGlnGluMetLysAsnMetSerArg 380  
 Db 1322 ATTAACCTTTCATGAGATTAATCCAACTTCACTTGAAGAAATGAAACATGAGCAGA 1381  
 QY 381 GlnGlyTyrValAlaHisLeuArgArgLysSerSerGlyPheSerArgGlyValAsnIle 400  
 Db 1382 CAGGAATAGTTGCACACTTGAGAAAGAAAGCAGCGGGTTCCTAGAGGCTTCACATA 1441  
 QY 401 TyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgVal 420  
 Db 1442 TACAGAGGGGTCAAGAGCATCAACATGGAATGGCAAGCAGATGAGCAGAGT 1501  
 QY 421 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheSerThrGlnGlnGluAlaAlaGluAla 440  
 Db 1502 GCTGGGAACAAGACCTTACCTTGAGAGCTTGCACACCCACAGAGAGAGCAGCAAGCA 1561  
 QY 441 TyrAspValAlaAlaIleLysPheArgGlyAlaAsnAlaValThrAsnProHisLeuSer 460  
 Db 1562 TACGATGAGCGGATCAAAATTTGGCGCGAAATGCAATGACATCAAACTTGAACATTCA 1621  
 QY 461 ArgTyrAspValGluArgIleMetLysSerSerAsnLeuLeuAlaGlyGluLeuAlaArg 480  
 Db 1622 AGATACAGATGGAGAGATCAATGCGCAGTACCAATCTCGTGGGAGCTTCAAGG 1681  
 QY 481 ArgLysLysAspAsnProArgAsnLysAspIleAspTyrAsnLysSerValAlaThr 500  
 Db 1682 CGTAAAGAAATACATCTTGAACAAGAGCATGACTTACAAACAGCTGATGATACA 1741  
 QY 501 SerValAsnAsnGluGlnThrValGlnValGlnAlaGlyAsnAsnAsnGluAsnAsp 520  
 Db 1742 AGTGTGACAAATGAGAAAGGTTCAAGTCAAGCAGAAACACATATGATGAAAGCAG 1801  
 QY 521 SerGluTrpLysMetValLeuPheAsnHisProSerGlnGlnGlnAlaAsnGlyAsn 540  
 Db 1802 TCAGAGTGAAGATGATTTATTTAACACCCCTTCACACAGCAAGCAAGCAATGGCAAT 1861  
 QY 541 GlySerAspGlnLysIleMetLysCysGlyAsnTyrArgAsnSerLysPheSerMetAla 560  
 Db 1862 GGCAGTACCAAAATTAATGAACTGTGAAATTAAGAAACAGTCAATTTCTATGGCC 1921  
 QY 561 LeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGlu 580  
 Db 1922 CTACAAATCTTATTTGGATGATTCGTGGGTTCGGGAGAGCATATATAGTCGACGAG 1981  
 QY 581 SerSerLysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSer 600  
 Db 1982 TTAGCAAAATTTGGAATCTTTTCAACACGCTCATCGTGTGACAACTTAAAGAGC 2041  
 QY 601 SerArgGluLysSerProGluLysArgLysProSerLeuLeuPheProMetProPheMet 620  
 Db 2042 TCAGAGAGCTAGTCTGAGAAAGGGGCTTCGCTGCTTCTTCCCAAGCTTCCAAATG 2101  
 QY 621 GluThrLysIleValAlaProIleGlyThrSerValThrSerTrpLeuProSerProThr 640  
 Db 2102 GAAACCAAGATTTGAACCCCATTTGATCAAGCTTACCTTGGCTACCTCAACAGC 2161  
 QY 641 ValGlnMetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrp 660  
 Db 2162 GTTCAAAATGAGGCTTCTCCGCTATCTTTGTGCTCACTGCAAGTTTGTGCTTCTGG 2221  
 QY 661 ThrAspThr 663





```

/ Sequence 18711, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313) B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 18711
/ LENGTH: 1378
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3109-011-F8_FLI
/ US-10-425-114-18711

```

```

Alignment Scores:
Pred. No.: 1,378-178 Length: 1378
Score: 1744.00 Matches: 331
Percent Similarity: 93.26% Conservative: 15
Best Local Similarity: 89.22% Mismatches: 13
Query Match: 49.80% Indels: 12
DB: 12 Gaps: 4

```

US-10-024-632-2 (1-663) x US-10-425-114-18711 (1-1378)

```

QY 1 MetLysArgIleAsnGlnSerAsnGlnThrAspAspGlyAsnAsnHisAsnTrpLeuGly 20
DB 283 ATGAGCGCATGATAGAGTAAACAACCGATGATGGAACCAATCATACCTGATGGG 342
QY 21 PheSerLeuSerProHisMetLysMetGluAlaTrpSerAlaIleThr----- 36
DB 343 TTCTCTCTCTCGGCCACATGAAATGAGGTACTCTCGGCCACCTGTTCCGACAC 402
QY 37 ---ValProThrThrPheThrMetSerProSerGlnSerHisLeuSerAsnPheGlyMet 55
DB 403 AAGTCTCCGACCAACCTTCTCAATGTCCT-----TTCACATGTCACATCCGGAATG 456
QY 56 CysTrpGlyValGlyGluAsnGlnPheHisSerProLeuThrValMetProLeuLys 75
DB 457 TGTAAAGTGTGGGAAATGATGTAATCTTCATCTCCCTTACCGTTATGCTCTCAAG 516
QY 76 SerAspGlySerLeuGlnLeuGlnAlaLeuLysArgSerGlnThrGlnValMetVal 95
DB 517 TCTGAGGGGTCACTTGATCTTGGAGCTCTCAATGATCAACAACCAAGTATGCTG 576
QY 96 ProThrSerSerProLysLeuGlnLysPheLeuGlnGlyAlaThrMetGlyThrHisGlu 115
DB 577 CCAACTTCGCTCTCCGAAATGGAGACTCTTAAAGGTGACACTATGGAACCTCAACCA 636
QY 116 TyrGlnSerHisGlnArgGlyLeuSerLeuAspSerIleTyrTyrAsnSerGlnAsnAla 135
DB 637 TATGGAGACCAACGAGAGAGGTGGAGCTTGAACAGATCTATATATCAACCAAAACCA 696
QY 136 GlnAlaGlnProAsnArgAspLeuLeuSerGlnProPheArgGlnGlnGlyHisMetSer 155
DB 697 GAGGCTCAACCAACGAAACCTTCTTCAATCCCTTGAAGCAACAAGGAGTGTGAT 756
QY 156 ValGlnThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeu 175
DB 757 GTCGAACACACACCTTATCTGCTGATTTGCTTGCAGGTTTATATCAAGCACCGTCG 816
QY 176 GlnGlnGlnThrThrHisGlnThrHisValSerAspCysSerSerLeuMetProGlnMet 195
DB 817 GAGGAGAACCAACAAAGGAAACGACGTT-----TCAGTATGCTCTCAATG 864

```

```

QY 196 Thr---GlnGlyLeuLysAsnTrpValAlaProThrArgGlnPheSerThrHisGlnGln 214
DB 865 ACAGAGAGAGGTTTGGCAAACTGGGTAGCTCCAAACAGAGGAAATTCATCATCTACAGCAG 924
QY 215 ValLeuGlnGlnGlnMetAsnCysGlyMetGlyAsnGlnLysArgAsnGlyValSerLeuGly 234
DB 925 ATTCTGGAGCAACCAATGAACTGTGGCATTTGGATGATAGAGAACTGGGCTATCTGTTGA 984
QY 235 SerValGlyCysGlyGlnLeuGlnSerLeuSerLeuSerMetSerProGlySerGlnSer 254
DB 985 ACTGGGGGTGGAGATTGCAATCTCAAGCTTACTCTGATGAGTCTCGGCTTCAGTCT 1044
QY 255 SerCysValThrAlaProSerGlyThrAspSerValAlaValAspAlaLysArgGly 274
DB 1045 AGTTGTGACCTGCTCTCTTCTGAAACAGATTTCTGCTGTGATGCAAGAAAGAGAGG 1104
QY 275 HisAlaLysLeuGlnGlnLysGlnProValHisArgLysSerIleAspThrPheGlyGln 294
DB 1105 CATGCTAAACTTGGTCAGAACACACCTGTGCATGAAATCTATTGACACATTTGGGCAA 1164
QY 295 ArgThrSerGlnTyrArgGlyValThrArgHisArgTrpThrArgTyrGlnAlaHis 314
DB 1165 AGACGTCGCGATATAGAGCGCTCACAGGCATGATGATGATGATGATGATGATGATGAT 1224
QY 315 LeuTrpAspAsnSerCysLysLysGlnGlyGlnThrArgLysGlyArgGlnValTyrLeu 334
DB 1225 TTGTGGATTAATAGTTGCAAGAGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1284
QY 335 GlyGlyTyrAspMetGlnGlnLysAlaAlaArgAlaTyrAspLeuAlaLysLysTyr 354
DB 1285 GAGGTATGATATGAGAGAAAGAAAGCTGCAAGACCTTATGATCTCGGCTCTTAAGTAC 1344
QY 355 TrpGlyProSerThrHisIleAsnPheSerIle 365
DB 1345 TGGGAGCTTCAACGACATTAACCTTTGATA 1377

```

## RESULT 4

US-10-024-632-3

```

/ Sequence 3, Application US/10024632
/ Publication No. US20020170093A1
/ GENERAL INFORMATION:
/ APPLICANT: Monsanto Technology LLC
/ APPLICANT: He, Steve S.
/ APPLICANT: Dolson, Stanton B.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION
/ FILE REFERENCE: 38-21(51857) B
/ CURRENT APPLICATION NUMBER: US/10/024,632
/ PRIOR APPLICATION NUMBER: US 60/257,896
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3323
/ LENGTH: 2323
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (25)..(2022)
/ OTHER INFORMATION:
/ US-10-024-632-3

```

```

Alignment Scores:
Pred. No.: 4,246-175 Length: 2323
Score: 1715.00 Matches: 378
Percent Similarity: 65.86% Conservative: 83
Best Local Similarity: 54.00% Mismatches: 165
Query Match: 48.97% Indels: 74
DB: 13 Gaps: 20

```

US-10-024-632-2 (1-663) x US-10-024-632-3 (1-2323)

QY 1 MetLysArgGlyLeuSerGlnSerAsnSerHisThrAspArgGlyAsnAsnHisAsnThrLeuGly 20  
 DB 25 ATGAAGAGATGTGAAAATGATGACATGCTGCTTAATATCAAAACAAATGTTGGCT 84  
 QY 21 PheSerLeuSerProHisMet-----LysMetGluAlaThr 32  
 DB 85 TTCTCACTCTCTCCCAATGCAATAATAGAGTTTCTTCACACTGCACACCTTCCCT 144  
 QY 33 SerAlaAlaThrValProThrThrPheTyMetSerProSerGlnSerHisLeuSerAsn 52  
 DB 145 GCTGCTGAAGTGGTCTCTCAAGCTTTTACCAACACACTGCTCCA-----CTTAGTAC 198  
 QY 53 PheGlyMetCysTyGlyVal---GlyGluAsnGlyAsnPheHisSerProLeuThrVal 71  
 DB 199 TATGGTTTCTACTATGAGCTTGAAGCTGAAAATGTTGATTTGTATTCAGCTTTGCCAATC 258  
 QY 72 MetProLeuLysSerAspGlySerLeuCysIleLeuGluAlaLeuLysArgSerGlnThr 91  
 DB 259 ATGCCCTCAAAATCTGATGCTCTCTCTATGATTTGAAAACCTTTAGCGAGTCACAGCA 318  
 QY 92 GlnValMetValProThrSerSerProLysLeuGluAspPheLeuGlyValAlaThrMet 111  
 DB 319 CAAGCAATGGCTACTACTTCAACACCAAACTGGAGAACTTTTGAAGTGGGAGAACCAATG 378  
 QY 112 GlyThr---HisGluTyR-----GlySerHisGluArgGlyLeuSerLeuAspSer 127  
 DB 379 GGGACCCCTCATCATGACCAATGATAGTGCACAGAAACAATGCTCTGAGCTTAAACAT 438  
 QY 128 IleTyTyTrpAsnSerGlnAsnAlaGluAlaGlnProAsnHisArgAspLeuSerGlnPro 147  
 DB 439 GTTTTATAC---ATCCAACTGACGCGCTGACCCCAATATATACAAACCTTACCAAAAC 495  
 QY 148 PheArgGln-----GlnGlyHisMetSerValGlnThrHisProGlyTyR 162  
 DB 496 CATGTTCAACACATTCAGCAACCAACCAACCAACCAACAGAGCTTCAAGCATATATAC 555  
 QY 163 SerGlyLeuAlaCysHisGlyLeuTyGlnAlaProLeuGlnGluGlnThrHisGly 182  
 DB 556 TCTACCTTGAAGAACCATGATATGATATTA-----GAAAGGTCAAAGCA 600  
 QY 183 ThrHisValSerAspCysSerSerLeuMetProGln----- 194  
 DB 601 AGCCAACTTCTGCAACCAACATCTTCATCTTCAAAACATGGGTGATGATGCCGTT 660  
 QY 195 MetThrGluGlyLeuLysAsnThrValAlaProThrArgGluPheSer-----Thr 211  
 DB 661 CCGTTCCTGGCCCTCAAGAGTTGG-----GAAAGTGAAGAACTTCCAAAGCTACCAATGCA 714  
 QY 212 HisGlnGln-----ValLeuGlnGlnGlnMetAsnCysGlyMetGlyAsnGluArg 228  
 DB 715 CATAGCAAAAGATGATTTGCTCATGTGAGAAAATGCTGTGATCAATGAGG----- 768  
 QY 229 AsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeuSerMet 248  
 DB 769 -----TCCATTGATCAATGAGCTTATGCTGACCTTGCATGCTGAGCTTGCATG 819  
 QY 249 SerProGlySerGlnSerSerCysValThr-----AlaProSerGlyThr 263  
 DB 820 AGTCTCACTCTCACTGTCAGTGCAGAGTTCACAGTTTCTCACCGTCTTCACTGCTGTGTT 879  
 QY 264 AspSerValAlaValAspAlaLysLeuArgGlyHisAlaLysLeuGlnGlnLysGlnPro 283  
 DB 880 GATTCTGTCGCAATGATACATAAGAAAAGGGGCTGAAAAGCTTGCACAGAACAAAT 939  
 QY 284 ValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyArgGlyValThr 303  
 DB 940 GTTCATAGGAAGTCCATTCATGACCTTTGCAAAAGAACCTCCGATATAGAGAGCTAAC 999  
 QY 304 ArgHisArgGlyTrpThrGlyArgTyGlyAlaHisLeuThrPheAsnSerCysLysLysGln 323  
 DB 1000 AGGCAATAGGTGACCTGGAGATATAGACTATCTTTGGGCAACACAGCTGCAAGAAAGAG 1069  
 QY 324 GlyGlnThrArgLysGlyValArgGlnValTyLeuGlyGlyTyTrpAspMetGluGluValAla 343

DB 1060 GGCAAGAGCGAGAAAGAAAGACAAAGTTTATCTAGGGGGTTATGATATGAAAGAAAGCT 1119  
 QY 344 AlaArgAlaTyTrpAspLeuAlaAlaLeuLysTyTyTrpGlyProSerThrHisIleAsnPhe 363  
 DB 1120 GCGAGAGCTTATGATCTTACGGGCACTCAAGTATTTGGGAGCCCTCCACTACATTAACCTTT 1179  
 QY 364 SerIleGluAsnTyGlnValGlnLeuGluGlyMetLysAsnMetSerArgGlnGlyTyR 383  
 DB 1180 CTTTGAATAATTATCAAAATGAATGAGAAATAGAGAACATGACTAGACAAAGATAT 1239  
 QY 384 ValAlaHisLeuArgThrArgLysSerSerGlyPheSerArgGlyValaSerIleTyArgGly 403  
 DB 1240 GTTGCTATTGGAAGAAAAGACAGGCGATTTCTCAAGAGGGGCTTCCATGTCAGAGGA 1299  
 QY 404 ValThrArgHisHisGlnHisGlyArgTyGlnAlaArgIleGlyArgValAlaGlyAsn 423  
 DB 1300 GTAACAAAGACACCAACCAATGGAAGGTGGCAAGCTGAAATGGTAGAGTGGCTGGAAC 1359  
 QY 424 LysAspLeuTyLeuGlyTyTrpPheSerThrGlnGlnGluAlaAlaGlyValAspVal 443  
 DB 1360 AAAGATCTATATCTTGAACCTTTAGTACCAAGAGAGACAGCTGAGCTTATGATAT 1419  
 QY 444 AlaAlaIleLysPheArgGlyAlaAsnAlaValThrAsnPheAspIleSerArgTyAsp 463  
 DB 1420 GCTGCTATTAATTTCCAGAGAGCGAATGCTGTAAACAACTTTCATCACAAGATATGAT 1479  
 QY 464 ValGluArgIleMetAlaSerSerAsnLeuAlaGlyGluLeuAlaArgArgLys 483  
 DB 1480 GTGAGAAAATCATGCGAAGGACGACACCTCTTACGAGTACGCTAGGCGGAGCA 1539  
 QY 484 AspAsnSerProCysAsnLysAspIleAspTyTrpAsnLysValValThrSerVal--- 502  
 DB 1540 GAGAGCGCAATGAATCTCACTGATGATTAATAATGACATAACCTTGCATATGAG 1599  
 QY 503 AsnAsnGlnGluThrValGlnValAlaGlyAsnAsnAsnAsnGluAsnAspSerGlu 522  
 DB 1600 GACACTCAAGAGAGCTATCTAATGACCAAGAGAGCTGTGAGCAAAATGAT---CAG 1656  
 QY 523 TrpLysMetValIlePheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySer 542  
 DB 1657 TGGAGAGTGTCTCTAC---CAATCCTGAGCACTTGGAGCAATCCCAACAAT 1713  
 QY 543 AspGlnLysIleMetAsnCysGlyAsnTyTrpAsnSerAlaPheSerMetAlaLeuGln 562  
 DB 1714 CAG-----AGTACAGAACTAACCCAGTCTTGCAGTGGCTTTTGAC 1755  
 QY 563 AspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSer 582  
 DB 1756 AACATGTTT-----CATCAGAAAGTGAAGATATCAAGT 1788  
 QY 583 LysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerArg 602  
 DB 1788 AAGCGAGAGCCATGTGTCAAATCTTCTTCAATGGCCACAAGTTTGAAGCGCTCAAGA 1848  
 QY 603 GlnAlaSerProGlnLysArgGlyProSerLeuLeuPheProMetProPrometGlnThr 622  
 DB 1849 GAAGGTAGCCCTGATAGACAAAGCTTGCACATGCTCTGGAATCCCTTCAACTGCATCA 1908  
 QY 623 LysIleValAsnProIleGlyThrSerValThrSerTrpLeuProSerProThrValGln 642  
 DB 1909 AAACATTTGGCTACTAATCCAAATACATGATTTTGGAGCCCTTCAACCCCATTTGAGG 1968  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerTrpHisSer 662  
 DB 1969 -----CAGCACTTACTTGTGCTCAAAATGCCAGTTTTCAGAGCTTGAAGAT 2016

RESULT 5  
 US-10-024-632-5  
 ; Sequence 5, Application US/10024632  
 ; Publication No. US20020170093A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Monsanto Technology LLC

APPLICANT: He, Steve S.  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
 FILE REFERENCE: 38-21(51837)B  
 CURRENT FILING DATE: 2001-12-19  
 PRIOR APPLICATION NUMBER: US 60/257,896  
 NUMBER OF SEQ ID NOS: 33  
 SOFTWARE: Patent version 3.1  
 SEQ ID NO: 5  
 LENGTH: 1926  
 TYPE: DNA  
 ORGANISM: Oryza sativa  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(1926)  
 OTHER INFORMATION:  
 US-10-024-632-5

Alignment Scores:  
 Pred. No.: 5.01e-142 Length: 1926  
 Score: 1408.50 Matches: 340  
 Percent Similarity: 55.05% Conservative: 63  
 Best Local Similarity: 46.45% Mismatches: 152  
 Query Match: 40.22% Indels: 177  
 Gaps: 26

US-10-024-632-2 (1-663) x US-10-024-632-5 (1-1926)

QY 13 GYAAEAHMSASNTPLLEUGLYPHESELEUSEPROHISMETLYSMETGLUALATHR 32  
 DB 13 GGGCGACAGAGAACTGGTTAGGCTTCGCTCCCGACATGCCGCCATGAGAGTG 72  
 QY 33 SerAlaAlaThrValProThrThr 40  
 DB 73 CGGTCTCTCTGACCCATCGACTGCTCATCATCATCATCATCATCATCATCTGCT 132  
 QY 41 -----Phe 41  
 DB 133 GGTGCTGCTGCTGCCGAGCATGTCCTCTCCCGACAGCGACAGCTGCACTTC 192  
 QY 42 TyrMetSerProSerGlnSerHisLeu-----SerAsnPhgGly 54  
 DB 193 CTCTTCTCTCCCTCTGACAGACAGATGCTGCTCTTCCCTGACTACTACTACTGCTGGC 252  
 QY 55 MetCysTyrGlyValGlyGluAsn-----GlyAsnPhen:SerProLeuThrValMet 72  
 DB 255 GGGCGCTACGAGAGGAGACAGACCGCGCGCTCTACTCTGACACTTCCCTGTCATG 312  
 QY 73 ProLeuYSerAspGlySerLeuGlySerLeuGlyAlaLeuLysArgSerGlnThrGln 92  
 DB 313 CTTATCAAGTCCGATGGCTCCCTCTGACATCATGGAAGC-----351  
 QY 93 ValMetValProThrSerSerProLysLeuGluAsnPhenGlyAlaThrMetGly 112  
 DB 352 ---ArgATGCCCTGCTGATGCCAAAGCTCGAGACTCTTGGGGTGGCAATGGCACT 408  
 QY 113 ThrHisGlyTyrGlySerHisGlyLeuArgGlyLeuSerLeuAspSerIleTyrTyrAsnSer 132  
 DB 409 GGGCAT-----GACCCGCGCCACTACTACTATAGC 435  
 QY 133 GlnAlaAlaGluAlaGlnProAsnArgAspLeuLeuSerGlnProPhaArgGlnGly 152  
 DB 436 CAGGCGCAAGAGAGAGAT-----GCAAGCAGGCGCGCTTACAGAC 480  
 QY 153 HisMetSerValGlnThrHisProTyrTyrSerGlyLeuAlaCysHisGlyLeuTyrGln 172  
 DB 481 CACCACTAGTC-----CCCTAACAC-----TACAG 507  
 QY 173 AlaProLeuGluGluThrThrHisGlyThrHisValSerAspCysSerSerLeuMet 192

DB 508 CATTAGCAGAGAGAGATGCTGCAAGAGCGCGGCGCCATGAGAGAGCAATG 567  
 QY 193 ProGlnMetThrGluGlyLeuLysAsnTrpValAlaProThrArgGlyPheSerThrHis 212  
 DB 568 -----GGGCGGCGCAAGAACTTC-----585  
 QY 213 GlnGlnValLeuGluGlnGlnMetAsnGlyMetGlyAsnGluArgAsnGlyValSer 232  
 DB 586 -----CTGTCACAGCTACGCGCGCTGCTACGCAAGCAACAGAGATGCCG-----630  
 QY 223 LeuGlySerValGlyCysGlyGlyLeuGlnSerLeuSerLeuSerMetSerProGlySer 252  
 DB 631 -----CAGCGCTCAGACTCTCTCATGAGCCAGGCTCC 663  
 QY 253 Gln---SerSerCysValThr---AlaProSerGlyThrAspSerValAlaValAspAla 270  
 DB 664 CAGTCCAGCAGCTGCTGCTACGTGCACTCCAGCAGCATCAGCAGATGCGGTGGTCT 723  
 QY 270 -----270  
 DB 724 GCAGCTGCTGCTGCTGATGCGCAGGAGCAAGTAATGACGTGGCGAGCAGCT 783  
 QY 271 -----LysLysArgGlyHisAlaLysLeuGlyGlnLysGlnProValHisArgPheSer 288  
 DB 784 GTGCGAAGAGAGAGGCGACCGGAAAGGCGCCAAAGCAGCTTTCACCGAAGTCC 843  
 QY 289 IleAspThrPheGlyGlnArgThrSerGlnTyrArgGlyValAlaThrArgHisArgTyrThr 308  
 DB 844 ATTGACAGCTTGGGACAGAGACATCGCATATAGGGCGCTCACCGACAGCAGGCGACT 903  
 QY 309 GlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysLysGluGlyGlnThrArgHis 328  
 DB 904 GGAAGATATGAAGCCCACTCGGATACAGTTGCAAAAGAGATGGACAGACAAAGAG 963  
 QY 329 GlyArgGlnValTyrLeuGlyGlyTyrAspMetGluGlyValAlaAlaArgAlaTyrAsp 348  
 DB 964 GGAAGCAAGTATCTAGTGGTGTATGACTGAAGATAAGCTGCAGAGGCTTATGAT 1023  
 QY 349 LeuAlaAlaLeuLysTyrTyrGlyProSerThrHisIleAsnPhenSerIleGluAsnTyr 368  
 DB 1024 CTGGCTGCGCTGAATACTGGGGGCTATCTAGCATTAATAATTCCTCGTGAATACTAC 1083  
 QY 369 GlnValGlnLeuGlnGlnMetLysAsnMetSerArgGlnGlyTyrValAlaHisLeuArg 388  
 DB 1084 CAGATGAGATCGAGAGATGGAAGAGATGAAGAAGCAAGATATGTTGGCAGCTTGAGA 1143  
 QY 389 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValAlaThrArgHis 408  
 DB 1144 AGGAGAGAGAGGGGTTCTCTGCGGTCTTCATCTACCGGGAGTAACAAGGATCAC 1203  
 QY 409 GlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 428  
 DB 1204 CAGATGAGATGAGAGAGCTGGATGGCAGGCTTGTGGCAACAAGACTTGATCTC 1263  
 QY 429 GlyThrPheSerThrGlnGluGlnAlaAlaGluAlaTyrAspValAlaAlaIleLysPhe 448  
 DB 1264 GGCATTTCAGCATCTCAAGAAAGAGAGAGAGGATGACATTCGCTGCCCTCAAGTTC 1323  
 QY 449 ArgGlyAlaAsnAlaValThrAsnPhenAspIleSerArgTyrAspValGluArgIleMet 468  
 DB 1324 CGTGGCTGAAGCGGTGCGAAGCTTTGACATCACAAGGTAGAGCTGGACAAAGATCATG 1383  
 QY 469 AlaSerSerAsnLeuLeuAlaGlyLeuLeuAlaArgArgLys-----483  
 DB 1384 GAGAGCAGCTGCTGCTGCTGCTGAGCAGCGCGTAAGTAAAGCGATCGAGCAGAGCG 1443  
 QY 484 ---AspAsnAspProArgAsnLysAspIleAspTyrAsnLysSerValAlaThrSerVal 502  
 DB 1444 CCGGACCATGTGCAATAGCGCGAGCTC-----GGTGG 1479  
 QY 503 AsnAsnGluGluThrValGlnValGlnAlaGlyAsnAsnAsnAsnGluAsnAspSerGlu 522  
 DB 1480 ACCGAGAGAGAGCGCTGCTACTGTCACGGC-----ACCGAC 1518



QY 345 rgaAaTyrAaPleuAaIaIaLeuYrTyrGlyProSerThrHisIleAsnSerI 365  
 Db 1055 GGGCGATGATGATCTTGGCGCTCAAGTACTGGGGGCGCTCCACGACATCAACTCCCGT 1114  
 QY 365 IeGIuAenTyrGlnValGlnLeuGluMetIleAsnMetSerArgIleGlnTyrValA 365  
 Db 1115 TGAAGAACTACAG 1174  
 QY 385 IahIleuAaArgIleSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValT 405  
 Db 1175 CTCACTCAGAAAGAAAG 1234  
 QY 405 hrrAghIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 425  
 Db 1235 CCAAGCATCATCAG 1294  
 QY 425 sPleuTyrLeuGlyThrPhe----- 431  
 Db 1295 ACCCTTACTTGGAGAGATTCATCGCGCTTTCGCGCGCGCGCGCGCGCGCGCATG 1354  
 QY 432 --SerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 451  
 Db 1355 CCGGACAG 1414  
 QY 451 IahsnaIaValThrAsnPheAspIleSerArgTyrAspValGluArgIleMetAlaSerS 471  
 Db 1415 TCACAGCGCTCAGCACTTCGACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1474  
 QY 471 exaIleuLeuAaIaGlyIleuAaIaArgArgIle---LysAspAsnAspProArgAsnI 490  
 Db 1475 GCACGCTCTCCCGGGAG 1533  
 QY 490 ysaPleIleAspTyrAsnLysSerValIleThrSerValAsnAsnGlnIleThrValGln 510  
 Db 1534 -----GCTTGG 1567  
 QY 510 aIaGlnaIaGlyAaAsnAaAsnAaGlnAaAspSerGluTyrPheMetValLeu----- 527  
 Db 1568 TCGAGCGCGCG-----AACGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1612  
 QY 528 -----PheAsnHisProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 545  
 Db 1613 CGCTGCAGCGCGCGAG 1671  
 QY 545 ysrIleMetIleAsnGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuI 565  
 Db 1672 -----CACCATGACCTTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1720  
 QY 565 IeGly---IleAspSerValGlySerGlyGlnHisMetLeuAspGlnSerSerIysI 584  
 Db 1721 TGTTCGACCGTGCAGCG 1755  
 QY 584 IeGlyThrHisPheSerAsnThrSerSerIleValThrSerLeuSerSerArgGlnA 604  
 Db 1756 --GCGCGCGAGCATGTGATGCGCGCG-----ACAGAGCTGGGCAACTCCGGAGAGC 1804  
 QY 604 IaserProGluLysArgGlyPro-----SerLeuLeuP 615  
 Db 1805 AGAGCGCTGAC--AGAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCTTCCTCCACAGCGGT 1861  
 QY 615 hePheMetProPheMetGluThrIleValAsnProIleGlyThrSerValThrSerT 635  
 Db 1862 TCGCGAAGCG 1915  
 QY 635 rPleu---ProSerProThrValGlnMetArgProSerProAla---IleSerLeuSerH 653  
 Db 1916 GGGCGCTGCGCTCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1975  
 QY 653 IseuProValPheAlaSerTyrThrAsp 662  
 Db 1976 ACCTGCCAATGTTCCGCGCGGTGAGACGAC 2004

US-10-059-911-1  
 / Sequence 1, Application US/10059911  
 / Publication No. US20030159180A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Fischer, Robert L.  
 / APPLICANT: Mizukami, Yukiko  
 / APPLICANT: The Regents of the University of California  
 / TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 / FILE REFERENCE: 023070-090730US  
 / CURRENT APPLICATION NUMBER: US/10/059, 911  
 / CURRENT FILING DATE: 2002-06-17  
 / NUMBER OF SEQ ID NOS: 42  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 1  
 / LENGTH: 2148  
 / TYPE: DNA  
 / ORGANISM: Arabidopsis thaliana  
 / FEATURE:  
 / OTHER INFORMATION: AINTEGUMENTA (ANT) CDNA  
 / FEATURE:  
 / NAME/KEY: CDS  
 / LOCATION: (269)..(1936)  
 / OTHER INFORMATION: Arabidopsis AINTEGUMENTA (ANT)  
 US-10-059-911-1  
 Alignment Scores:  
 Score: 3,116-133 Length: 2148  
 Percent Similarity: 1328.00 Matches: 312  
 Best Local Similarity: 52.84% Conservative: 69  
 Query Match: 43.27% Mismatches: 116  
 DB: 37.92% Indels: 224  
 Gaps: 21  
 US-10-024-632-2 (1-663) x US-10-059-911-1 (1-2148)  
 QY 1 MetIleArgIleAsnGluSerAsnAsnThrAspAspGlyAsnAsnHisAsnTyrPleuGly 20  
 Db 269 ATGAAGTCTTTTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 328  
 QY 21 PheSerLeuSerProHis--MetIleMet----- 29  
 Db 329 TCTCATTTCTCTCAAAATGATGATAAATGGAGAGTAGAGAGAGTAGAGAGAGAGAGAGAG 368  
 QY 30 -----GluAlaThrSerAlaIaIaThrValProThrThrPheTyrMet 43  
 Db 389 TCATCTTCACTTCTTCACTGAGTCAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 448  
 QY 44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCysTyrGlyValGlyGluAsnGly 63  
 Db 449 GGT-----GACACACATGACAACTTGGTGTGCTATGATGATTAACCAATGGA 499  
 QY 64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeu 83  
 Db 500 GGAATCTATTCACATGCTGTGATGCACTGATGATGATGATGATGATGATGATGATGATG 559  
 QY 84 GluAlaLeuLysAspSerGlnThrGlnValMetValProThrSerProLysPleuGly 103  
 Db 560 GAAGCTCTCAACAGATCTTCTCACTGCAATCAATCAATCAATCAATCAATCAATCAATCA 619  
 QY 104 AspPheLeuGlyGlyAlaIaIaThrMetGlyThrHisGlnTyrGlySerHisGlnArgGly 122  
 Db 620 GATTTCTTGGG-----ACCATCAACAACAACAAGTCAACAAGAAGCATG 667  
 QY 123 ---LeuSerLeuAspSerIleTyrTyrAsnSerIleAsnIaGlnIaGlnProAsnArg 141  
 Db 668 GATCTTAGCTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
 QY 142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlnHisMetSerValGln 157  
 Db 719 ACTACAACTTTCAAGAAGTCTTAGCTTCCCTCAACAACAGAAACCAT----- 766  
 QY 158 ThrHisProTyrTyrSerGlyLeuAlaCysHisIleLeuTyrGlnAlaProLeuGlnGlu 177









Db 1429 GAA----- 1431  
 QY 526 ValLeupheAsnHisProSerGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnLys 545  
 Db 1432 -----GACCAAGC 1440  
 QY 546 IleValAsnCysGlyAsnIYrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIle 565  
 Db 1441 GCTCTAAATGCT----- 1452  
 QY 566 GlyIleAspSerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerIleGly 585  
 Db 1452 ----- 1452  
 QY 586 ThrHisPheSerAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluAlaSer 605  
 Db 1453 -----GTTGTGGAAGTGTGTTCCAAAGAAAGAGTCAGT 1485  
 QY 606 ProGluAsnArgGlyProSerLeu-----LeuPheProMetProProMetGluThrLys 623  
 Db 1486 ACTCCCGAGAGACTCTGAGTTTCCGCGCATTTTCCGCTTCAAGTTATCAAAAG 1545  
 QY 624 IleVal-----AsnProIleGlyThrSerValThrSerIleuProSerProThrValGln 642  
 Db 1546 ATGTTGGATCAAAATATGGCGGAATATGAGTCTTGACATCAACCTTAATGCTGAG 1605  
 QY 643 MetArgProSerProAlaIleSerLeuSerHisLeuProValPheAlaSerIleProThrAsp 662  
 Db 1606 CTTAAG--ACCGTCGCTTACTTGTGCTCAGATGCGGTTTGGCTGCTGAGT 1662  
 QY 663 Thr 663  
 Db 1663 TCT 1665

RESULT 10  
 US-10-059-911-4  
 ; Sequence 4, Application US/10059911  
 ; Publication No. US20030159180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fischer, Robert L.  
 ; APPLICANT: Mizukami, Yukiko  
 ; APPLICANT: The Regents of the University of California  
 ; TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 ; FILE REFERENCE: 023070-090730US  
 ; CURRENT APPLICATION NUMBER: US/10/059, 911  
 ; NUMBER OF SEQ ID NOS: 42  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 1738  
 ; TYPE: DNA  
 ; ORGANISM: Brassica napus  
 ; FEATURE:  
 ; OTHER INFORMATION: Canola AINTEGUMENTA (ANT) ortholog partial cDNA  
 ; OTHER INFORMATION: including coding region  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(1647)  
 ; OTHER INFORMATION: Canola AINTEGUMENTA (ANT)  
 US-10-059-911-4

Alignment Scores:  
 Pred. No.: 4 226-131 1738  
 Score: 1307.00 Matches: 306  
 Percent Similarity: 54.52% Conservative: 78  
 Best Local Similarity: 43.50% Mismatches: 104  
 Query Match: 37.32% Indels: 218  
 DB: 14 Gaps: 25

US-10-024-632-2 (1-663) x US-10-059-911-4 (1-1738)

QY 9 AsnThrAspAspGlyAsnAsnHisAsnTrpLeuGlyPheSerLeuSerProHisMet--- 27

Db 16 CATTAATGATGATGATTAATACACTTAATTTGTTAGGTTCTGCTGTCTTCAATATGTTG 75  
 QY 28 -----LysMetGluAlaThrSerAlaAlaThrValProThrThrPhe 41  
 Db 76 AAATGGGTGGTGAAGAGCTCTTACTCATCTTCGCTGCTTCAAGTGAACCTCT-- 132  
 QY 42 TyrMetSerProSerGln-----SerHisLeuSerAsnPheGlyMetCysTyr 57  
 Db 133 ---TCTGTTCACACAGACTTGTGTTGGCCACACAGTACCACTATGAGTTGCTAC 189  
 QY 58 GlyValGly---GluAsnGlyAsnPheHisSerProLeuThrValMetProLeuLysSer 76  
 Db 190 GGTCTTAAGTACAGCTAGAGGAATGATTTCTGAATGCTGTGATGCTCCCTCAGATCT 249  
 QY 77 AspGlySerLeuCysIleLeuGluAlaLeuLysAspSer----- 89  
 Db 250 GACGGTTCTCTTGTATATGAGAGCTTCAACAGATTTCTCACTCGAATATATCATCAC 309  
 QY 90 GlnThrGlnValMetValProThrSerSerProLysLeuGluAspPheLeuGlyValAla 109  
 Db 310 CATAGTCAAGT-----TCATCTCCAAAGATGAAGATTTTC----- 345  
 QY 110 ThrMetGlyThrHisGluThrGly---SerHisGluAspGly-----LeuSerLeuAsp 126  
 Db 346 ---TTTGGACCCCATCATCAACACAGTCAACAAAGACCCAGATCTTAGTATGAT 402  
 QY 127 SerIleYrTyAsnSerGlnAsnAlaGluAlaGlnProAsnArgAspLeuSerGln 146  
 Db 403 AGTTATTTCTAATAATACCATCTCATCG-----CCAAACAAACACCACTTTC 453  
 QY 147 ProPheAsnGlnGlnGlyHisMetSerValGlnThrHisProTyrTySerGlyLeuAla 166  
 Db 454 GAGTTC-----TTTAGCTTCCCTCAAACTAGAAACACCAT-- 489  
 QY 167 CysHisGlyLeuTyGlnAlaProLeuGluGluGluThrThrHisValSer 186  
 Db 490 -----GAGAAAGAAAC-----AGAAACTACAGAGAT 516  
 QY 187 AspCysSerSerLeuMetProGlnMetThrGluGlyLeuLysAsnTrpValAlaProThr 206  
 Db 517 GAC-----CCTGTTGACACATGAA----- 537  
 QY 207 ArgGluPheSerThrHisGlnGlnValLeuGluGlnGlnMetAsnCysGlyMetCysLysn 226  
 Db 538 -----GAGGGCTCTTAAATGTAAGGCTATATGGG 567  
 QY 227 GluArgAsnGlyValSerLeuGlySerValGlyCysGlyGluLeuGlnSerLeuSerLeu 246  
 Db 568 GAATTCAA-----CAGTCACTGAGCTTG 591  
 QY 247 SerMetSerProGlySerGlnSerSerCysValThrAla----- 259  
 Db 592 TCCATAGCCCTGGGTCACAAATGACTGATCACTGCTCTCATCCACCAACCAACCA 651  
 QY 260 -----ProSer 261  
 Db 652 ACTCAAAACCAACGAGATCTCTGAAGCTTTGGTCAGACAAAGTCTGATTTGAGACA 711  
 QY 262 GlyThrAspSerValAlaValAspAlaLysLysArgGly---HisAlaLysLeuGlyGln 280  
 Db 712 ACACACATGGCGGCTCTGCTGCAAAAGAAAGAGAGCAAGAGATTTGCTGTGACAG 771  
 QY 281 LysGlnProValHisArgLysSerIleAspThrPheGlyGlnArgThrSerGlnTyArg 300  
 Db 772 AAACAGATTGTTCATAGAAATCTATTGACTTTTGACAAAGAACTTCGAAATCCGA 831  
 QY 301 GlyValThrArgHisArgTrpThrArgTyArgTyGluAlaHisIleuTrpAspAsnSerCys 320  
 Db 832 GGGCTTACACAGATATGATGAGTGTAGGTATGAAAGTCAATCTATGAGACAAATAGTTTC 891  
 QY 321 LysLeuGluGluGlnThrArgLysGlyArgGlnValTyLeuGlyGlyTyArgMetGlu 340

```

Db      892 AAGAGGAGGCTATGACAGAAAAGAGCAAGTTTATCTGGGGCTTATGATATGAG 951
Qy      341 GULYALALAAAGATATYRAAPLEUALAALALEULYETTPGLYProSerThrh 360
Db      952 GAGAAAGCTCTGAGCATATGATCTGTGCTGCACTAGACTGGGGTCCCTTACTATC 1011
Qy      361 ILAaNPheSerIleGluAenTYRGLNValGlnLeuGlnGluMetLysaMetSerArg 380
Db      1012 ACTAATTTCTCTGGGAAATATACGAAAGAGTTGATACATGAAGAACTGACTCA 1071
Qy      381 GlnGluTYRValAlaHisLeuArgLysSerSerGlyPheSerArgLysValSerIle 400
Db      1072 CAAGAATATGTGTCTGCTTGAAGAAAACAGTGGTTCTCTAGGGGTGCTTCATC 1131
Qy      401 TYRArgLYValAlaThrArgHisLeuGlnHisGlyArgThrGlnAlaArgIleGlyArgVal 420
Db      1132 TATAGAGGAGTCAACACATCACACAGATGAAAGGGAGAGCTGGATCGATGAGTC 1191
Qy      421 AlagLYaenLYsaAPLeuTYRLeuGlyThrPheSerThrGlnGlnAlaAlaGlnAla 440
Db      1192 GCTGGAACAAAGATCTCTACCTTGAACTTCGAACTCAAGAGAGCGCGGAAGCC 1251
Qy      441 TYRArgValAlaAlaIleLYsaPheArgLYAlaAsnAlaValThrAsnPheAPLeuSer 460
Db      1252 TATGATGTAGACATCACTCAAGTTCCGTGGCAAAACCGGTGACTTAACTTGAATPACA 1311
Qy      461 ArgTYRArgPValGluArgIleMetAlaSerSerAsnLeuAlaGlyGluLeuAlaArg 480
Db      1312 AGGATCAATGTGTGATCCGATTAAGCTAGTACACTCTCTGTCTGGAGAGAGCGCTCGA 1371
Qy      481 ArgLYsaPsaPsaPsaPsaPsaPsaPsaPsaPsaPsaPsaPsaPsaPsaPsaPsa 500
Db      1372 AGCAACAGC-----AACACATCGTGGTCCG 1398
Qy      501 SerValaAsnArgGln-----ThrValGlnValGlnAlaGlnAlaAsnAsn 517
Db      1399 AACATTAGCGAGAGAGAGCGCGCTTTAACCGCTGCGTGAACGCTGTTCCAAATAGAA 1458
Qy      518 GluAsnPheSerGluTyrIleMetValLeuPheAsnHisProSerGlnGlnGlnAla 537
Db      1459 GTG----- 1461
Qy      538 AsnGlyAsnGlySerAspGlnLysIleMetAsnGlyAsnTYRArgAsnSerAlaPhe 557
Db      1462 -----GGTAGCCCGAGAGGGGTTTGTAGTTT 1488
Qy      558 SerMetAlaLeuGlnAspLeuIleGlyLeaSerValGlySerGlyGlnHisAsnMet 577
Db      1488 ----- 1488
Qy      578 LeuAspGluSerSerLysIleGlyThrHisPheSerAsnThrSerSerLeuValThrSer 597
Db      1488 ----- 1488
Qy      598 LeuSerSerSerArgGluAlaSerProGlnLysArgGlyProSerLeuLeuPheProMet 617
Db      1489 -----CGAGAGATATTGGCGTGG 1506
Qy      618 ProPheMetGluThrLysIle-----ValaAsnProIleGlyThrSerValThrSerTyr 635
Db      1507 CCTCAAGTTGTGTCGAGAGATGTTCCGAGAGCAATGTGTCGGA---AAATAGAGTTCTTGG 1563
Qy      636 LeuProSerProThrValGlnMetArgProSerProAlaIleSerLeuSerHisLeuPro 655
Db      1564 ACTACAAACCTTAATGCTATGCTCAAG---ACCGTTCTCTTACTCTCCGAGATGCCG 1620
Qy      656 ValPheAlaSerTyrThrAspThr 663
Db      1621 GTTTTCGTCGCGGGCTGATTC 1644

```

RESULT 11  
US-10-059-911-30  
; Sequence 30, Application US/10059911

```

; Publication No. US20030159180A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; TITLE OF INVENTION: Methods for the University of California
; FILE REFERENCE: 023070-090730US
; CURRENT APPLICATION NUMBER: US/10/059,911
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1359
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: truncated or
; OTHER INFORMATION: mutated ANT cDNA transgene ANTDC1
; US-10-059-911-30

Alignment Scores:
Pred. No.: 2,64e-123 Length: 1359
Score: 1233.50 Matches: 273
Percent Similarity: 60.34% Conservative: 42
Best Local Similarity: 52.30% Mismatches: 82
Query Match: 35.22% Indels: 125
DB: 14 Gaps: 14

US-10-024-632-2 (1-663) x US-10-059-911-30 (1-1359)

Qy      1 MetLYsaGlyLeaAsnGluSerAsnAsnThrAspArgLYaAsnHisAsnTyrLeuGly 20
Db      1 ATGAAATCTTTTGTGATATGATATGATATATATATATACCAACAGCACTAATTGTAGG 60
Qy      21 PheSerLeuSerProHis---MetLYsaMet----- 29
Db      61 TTCTCATGTGCTTCAAAATGATGAAATGGAGGTAGAGAGGTAGAGAACTATTATAC 120
Qy      30 -----GluAlaThrSerAlaAlaThrValProThrThreTYrMet 43
Db      121 TCATCTTCACTTCTTCACTGCACTGCACTTCTTCTTCTGTTCCACCTCACTTGTGT 180
Qy      44 SerProSerGlnSerHisLeuSerAsnPheGlyMetCysTYRGLYValGlyGluAsnGly 63
Db      181 GGT-----GACAACTAGCAACTTGTGTTGTGATGAGATCTAACCCAAATGGA 231
Qy      64 AsnPheHisSerProLeuThrValMetProLeuLysSerAspGlySerLeuCysIleLeu 83
Db      232 GGAATCTATTTCCACTGTGTGTGAGCCACTCAAGTCTGATGTTCTTCTTCTTCTTATG 291
Qy      84 GluAlaLeuLYsaSerGlnThrGlnValMetValProThrSerSerProLYsaLeuGln 103
Db      292 GAAGCTCTCAACAGATCTTCACTGATGATGATGATGATGATGATGATGATGATGATG 351
Qy      104 AspPheLeuGlyGlyAlaThrMetGlyThrHisGlyTYRGLYSerHisGlyArgGly 122
Db      352 GATTTCTTGGG-----ACCATCAACAAACAAACAAAGTACAAAGAGCCATG 399
Qy      123 ---LeuSerLeuAspSerIleTYRLeuSerGlnAsnAlaGlnAlaProAsnArg 141
Db      400 GATCTTAGCTTAGATATGTTATTTCTTCAACACACACTCAAT-----GAGCCAAACAG 450
Qy      142 AspLeuLeuSerGlnProPhe-----ArgGlnGlnGlyHisMetSerValGln 157
Db      451 ACTACAAACTTTTCAAGAGTCTTTTACCTTCCCTCAACCAAGAAACAT----- 498
Qy      158 ThrHisProTYRtyrSerGlyLeuAlaCysHisGlyLeuTYRGLNAlaProLeuGlnGln 177
Db      499 -----GAGGAA 504
Qy      178 GluThrThrLYsaGluThrHisValSerAspCysSerSerLeuMetProGlnMetThrGln 197
Db      505 GAAACT-----AGAAATTACGGGAATGAC-----CTTAGTTTGAACAT 543

```

Page 14

	TITLE OF INVENTION: Methods for Altering Organ Mass in Plants
	FILE REFERENCE: 023070-090730US
	CURRENT APPLICATION NUMBER: US/10/059, 911
	CURRENT FILING DATE: 2002-06-17
	NUMBER OF SEQ ID NOS: 42
	SOFTWARE: Patentin Ver. 2.1
	SEQ ID NO 26
	LENGTH: 1422
	TYPE: DNA
	ORGANISM: Artificial Sequence
	FEATURE:
	OTHER INFORMATION: Description of Artificial Sequence:truncated or
	OTHER INFORMATION: mutated ANT CDNA transgene ANTND1
	US-10-059-911-26
	Alignment Scores:
	Pred. No.: 3.03e-122 Length: 1422
	Score: 1224.00 Matches: 284
	Percent Similarity: 53.46% Conservative: 56
	Best Local Similarity: 44.65% Mismatches: 92
	Query Match: 34.95% Indels: 204
	DB: 14 Gaps: 18
	US-10-024-632-2 (1-663) x US-10-059-911-26 (1-1422)
Qy	69 LeuThrValMetProLeuLysSerAspGlySerLeuValLeuGluAlaLeuLysArg 88
Db	1 ATGCTGTGATGCACTGACATGATCTGATGATCTCTTCTCTTAAGCAAGCTCTCAACAGA 60
Qy	89 SerGlnThrGlnValMetValProThrSerProLysLeuGluAspPheLeuGly 108
Db	61 TCTTCTCACTGCAATCAACATCAAGATTATCTCCAAAGGTGAGAGATTCTTTGGG--- 117
Qy	109 AlaThrMetGlyThrHisGluIuArgIysSerHisGluArgGly-----LeuSerLeuAsp 126
Db	118 -----ACCATCAACAACAACAAGTCAACAAGAAGCATGATCTTACCTTAGAT 168
Qy	127 SerLeuIuArgLysSerGlnHisAlaGluAlaGlnProAsnArgAspLeuLeuSerGln 146
Db	169 AGTTATTTCTTCAACACCACTCTCT-----GGGCCCAACAGCATCAAACTTTCAA 219
Qy	147 ProPhe-----ArgGlnGlnGlyHisMetSerValGlnThrHisProTyrTyr 162
Db	220 GAGTTCCTTAGCTTCCTCCCTCAACCAAGAAACAT----- 252
Qy	163 SerGlyLeuAlaCysHisGlyLeuTyrGlnAlaProLeuGluGluGluThrThyLysGlu 181
Db	253 -----GGGGAAGAAACT-----AGA 267
Qy	183 ThrHisValSerAspCysSerSerLeuMetProGlnMetThrGluGlyLeuLysAsnTrp 202
Db	268 AATTACGGGAATGAC-----CCTAGTTGACACATGGA----- 300
Qy	203 ValAlaProThrArgGluPheSerThrHisGlnGlnValLeuGluGlnGlnMetAsnCys 222
Db	301 -----GGGCTCTTTAAAGTGA 318
Qy	223 GlyMetGlyAsnGluArgGluAsnGlyValSerLeuGlySerValGlyCysGlyGluGluGln 244
Db	316 GGGGATATGGGGAAATTCA-----CAG 333
Qy	243 SerLeuSerLeuSerMetSerProGlySerGlnSerSerCysValThrAlaProSer--- 267
Db	340 TCACGTGACCTTATCATGACGACCTCTGGGTCAACAATCTAGCTGCATCATCTGCTTCACAC 399

```

QY 274 GLYHIS-----AlaIleuGlyGlnGlnProValHisArgLysSerIleAsp 290
DB 520 GGAACAAGAGATGTTTACTGTTGTGTGCAAAACAGATTGTTCAATAGAAAATCTATGAT 579
QY 291 ThrPheGlyGlnArgThiSerGlnTyArgGlyValIThrArgHisArgTyrThrGlyArg 310
DB 580 ACTTTGGACCAAGACCTTCTCAATACAGAGCGCTTACACACATAGATGACATCGTGA 639
QY 311 TYTGluAlaHisIleuTyrAspAsnSerCysIleLysGlnGlyGlnThrArgLysGlyArg 330
DB 640 TATGAAGCTCATCTATGAGACATATGTTCAAGAAAGAAAGGTCAACAGTAAAGAAAGA 699
QY 331 GlnValTyrLeuGlyGlyTyrAspMetGlnGlyValAlaAlaAlaGlyValAspLeuAla 350
DB 700 CAAGTTATCTGGAGAGTTATGATATGAGGAGAAAGCTGCTCCAGCATATGATCTTGT 759
QY 351 AlaLeuLysTyrTyrPLeuProSerThiHisIleAsnPheserIleGluAsnTyrGlnVal 370
DB 760 GACCTCAAGTACTGGGGGCTCCCTCTACTACACCAATTTCTCTGGGAGAAATTATCAGAA 819
QY 371 GlnLeuGlnGlnMetLysAsnMetSerArgGlnGlnTyrValAlaHisLeuArgLys 390
DB 820 GAGATTGAAGACATGAGAAACATGACTGACACAAAGATATGTTGACATTTGAGAGAGAG 879
QY 391 SerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValIThrArgHisIleGlnHis 410
DB 880 AGCAGTGGTTCTCTAGGGGCTCTCTCACTATACAGAGAGTCAACAGACATCACACAT 939
QY 411 GlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThr 430
DB 940 GGAAGGTGGCAAGCAAGATGGTGAAGTGGTGAACCAAGATCTCTTACTTGGAACT 999
QY 431 PheSerThiGlnGlnGlnAlaGlnAlaGlyValAspValAlaAlaIleLysPheArgGly 450
DB 1000 TTTCGAACCCAGAAACAGCTCCAGAGTTACGATGACAGCAATTAATTCGCGGC 1059
QY 451 AlaAsnAlaValIThrAsnPheAspIleSerArgTyrAspValGluArgIleMetAlaSer 470
DB 1060 ACAAAATGCTGTGACTAATCTTATATCAGAGGTACGATGATGATATGATGCTAGT 1119
QY 471 SerAsnLysLeuAlaGlyGlnLeuAlaArgArgLysAspAsnAspProArgAsnLys 490
DB 1120 AAACACCTCTTCTCGAGAGATTAGCGGCAAGAAC----- 1155
QY 491 AspIleAspTyrAsnLysSerValIThrSerValAsnAsnGlnGlnIThrValGlnVal 510
DB 1156 -----AACACAGCATTTGTC-----GTCAGCAATACTGA----- 1185
QY 511 GlnAlaGlyAsnAsnAsnGlnLysAspSerGluTyrLysMetValIleuPheAsnHis 530
DB 1185 ----- 1185
QY 531 ProSerGlnGlnGlnAlaAsnGlySerAspGlnLysIleMetAsnCysGly 550
DB 1186 -----GACCAACCGCTCTAAAGCT----- 1206
QY 551 AsnTyrArgAsnSerAlaPheSerMetAlaLeuGlnAspLeuIleGlyIleAspSerVal 570
DB 1206 ----- 1206
QY 571 GlySerGlyGlnHisAsnMetLeuAspGlnSerLysIleGlyThrHisPheSerAsn 590
DB 1206 ----- 1206
QY 591 ThrSerSerLeuValIThrSerLysSerSerArgGluAlaSerProGlnLysArgGly 610
DB 1207 -----GTTGCAAGGCTGCTCCACCAAGAAAGTCAAGTCCGCAAGAGATC 1254
QY 611 ProSerLeu-----LeuPheProMetProProMetGluThrLysIleVal---AspPro 627
DB 1255 TTGAGGTTTCCGGCGATTTTCGGCTTCCAGATTAATCAAAAGATGTTGGATCAAT 1314
QY 628 IleGlyThrSerValIThrSerTyrLeuProSerProThrValGlnMetArgProSerPro 647

```

---

```

DB 1315 ATGGCGGAAGATATGAGCTTGGACATCAACCTATGCTGAGCTTAAG---ACCGTC 1371
QY 648 AlaIleSerLeuSerHisIleuProValPheAlaSerThrPheAspThr 663
DB 1372 GCTCTTACTTTGCTCCATGATGCCGATTTTCGCTGCTTGGCGCTGATCT 1419

RESULT 13
US-10-425-114-11685
; Sequence 11685, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 11685
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701120037_F11
US-10-425-114-11685

Alignment Scores:
Pred. No.: 4,066-118 Length: 1231
Score: 1185.00 Matches: 237
Percent Similarity: 75.13% Conservative: 44
Best Local Similarity: 63.37% Mismatches: 63
Query Match: 33.84% Indels: 30
DB: 12 Gaps: 7

US-10-024-632-2 (1-663) x US-10-425-114-11685 (1-1231)
QY 290 AspThrPheGlyGlnArgThiSerGlnTyrArgGlyValIThrArgHisArgTyrThrGly 309
DB 1 GATACCTTTGGCAAAAGACATCTCAATATGAGGTGTAACAGCATAGGTGACAGCTGT 60
QY 310 ArgTyrGlnAlaHisIleuTyrAspAsnSerCysIleLysGlnGlyGlnThrArgLysGly 329
DB 61 AGATATGAAAGCTCATCTATGGGACAAACAGCTGCAAGAAAGAGGCGCCAAAGAAAGAGA 120
QY 330 ArgGlnValTyrLeuGlyGlyTyrAspMetGlnGlyValAlaAlaAlaArgAlaTyrAspLeu 349
DB 121 AGGCAAGTTTATCTAGGGGCTTATGATGAGAAAAAGAGCAAGCATATGATGATG 180
QY 350 AlaAlaLeuLysTyrTyrGlyProSerThiHisIleAsnPheserIleGluAsnTyrGln 369
DB 181 GCGGCACTCAAGTATGGGACCTCTCTCATATTAATTTCCATGGAATAATTATCA 240
QY 370 ValGlnLeuGlnGlnMetLysAsnMetSerArgGlnGlnTyrValAlaHisIleuArgArg 389
DB 241 AAGCAACTTGAAGAAATGAGACATGACGAGACAGAAATGTTGCTCATTTTCAAGG 300
QY 390 LysSerSerGlyPheSerArgGlyValAspIleTyrArgGlyValIThrArgHisIleGln 409
DB 301 AAAAGCAAGGATTTTCAAGAGGCGCTCCATGACAGAGAGATCAAGACACACCA 360
QY 410 HisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 429
DB 361 CATGGAAGTGGCAAGCTCGAATGGAAGGCTACCGGAAACAAAGATCTATATCTTGA 420
QY 430 ThrPheSerThiGlnGlnGlnAlaAlaGlnAlaTyrAspValAlaAlaIleLysPheArg 449
DB 421 ACCTTATGATCCAAAGAGAGAGAGAGCTGAACCTTATGACATGCTGCTATTAATTCAG 480

```

```

QY 450 GlyAlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGluArgIleMetAla 469
DB 481 GGAGTTAACTGCTGCACTACTTGGATATACAAAGATATACCTGGAAAAAATATATGAG 540
QY 470 SerSerAsnIleuLeuAlaGlyGluLeuAlaArgArgIleAsp---AsnAspProArg 488
DB 541 AGCAATTAACCTTCTTACAGAGTGAACAGCTTAAGCGGAAAGAGATGTGATGAGAACT 600
QY 489 AsnIleAspIleAspTyrAsnIleSerValIleThrSerValAsnAsnGluIleThrVal 508
DB 601 AGAAGCCAGGCTACCGCTTAACCAAAACCTTCAATATGACCACTCAAGAACCAT 660
QY 509 GluValGlnAlaGlyAsnAsnAsnGlnAsnAspSerGluTyrIleMetValIlePhe 528
DB 661 CTATATGCGAG-----AAAAGATGCAAAAACCAATGAGATGAGATGCTCACTTT 711
QY 529 AsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnIleMetAsn 548
DB 712 CCATGCCCCCAACAGCTTGATCAG-----AATCAAAAGATCGAGAGT 753
QY 549 CysGlyAsnTyrArgAsnSerAlaPheSerMetAlaGluIlePheIleGlyIleAsp 568
DB 754 TGT-----AGAACTCAGCCCTCTCAACGAGCTTAAGTAACTAGTTT----- 795
QY 569 SerValGlySerGlyGlnHisAsnMetLeuAspGluSerSerIleGlyThrHisPhe 588
DB 796 -----CGTACCAAGTTAGGAGACGAGCAACATGGAAACACACTTG 837
QY 589 SerAsnThrSerSerLeuValThrSerLeuSerSerArgGlnAlaSerProGluLys 608
DB 838 TCAATCTCTTCTCTCTGCTGACAGATTGAGTCAAGAGAGAGAGAGCCGAGTAA 897
QY 609 ArgGlyProSerLeuLeuPheProMetProMetGluThrIleValIleAsnProIle 628
DB 898 ACAACATGCAATGCTCTTTGGAATGCTTCAACAGTGTCAATATTTG----- 948
QY 629 GlyThrSerValIleThrSerIleuProSerProThrValGlnMetArgProSerProAla 648
DB 949 ---GCTAAGCTGATCTTGGATGATCTTCAATCTCAAG-----ACTGCG 993
QY 649 IleSerLeuSerHisLeuProValPheAlaSerIleThrAsp 662
DB 994 CTTTCTAGCTCAGATGCAATTTTGTCTGCTGGACAGAT 1035

```

## RESULT 14

```

US-10-424-599-123485
/ Sequence 123485, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 123485
/ LENGTH: 1949
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_82514C.1
US-10-424-599-123485

```

```

Alignment Scores: 2,39e-116 Length: 1949
Pred. No.: 1171,50 Matches: 270
Score: 56,58% Conservative: 44
Percent Similarity: 48,65% Mismatches: 67
Best Local Similarity:

```

```

Query Match: 33.45% Indels: 175
DB: 12 Gaps: 15
US-10-024-632-2 (1-663) x US-10-424-599-123485 (1-1949)
QY 256 CysValThr---AlaProSerGlyThrAspSerValAlaValAspAlaLysArgGly 274
DB 2 TGTGTTAAGTTACCTTCACTGCTATCACTAATCTGTGGCACTGATACCAAGAAAGGGGG 61
QY 275 HisAlaLysLeu-----GlyGlnLysGlnProValHisAlaGlySerIleAspThrPhe 292
DB 62 CTTCAATATGCTGACCAAGATACAGAGCAATATAGCTCATAGAAATCCATATGATCCTT 121
QY 293 GlyIleArgThrSerGlnTyrArgGlyValThrArgHisArgTyrPheGlyArgGlyGlu 312
DB 122 GGGCAAGAACCTTCAATATAGAGTGTATCAAGGCACTAGTGAGCTGATGATATATA 181
QY 313 AlaHisLeuTyrAspAsnSerCysIleLysGluGlyGlnThrArgGlyArgGlyVal 332
DB 182 GCTCATCTATGGACACACAGCTGCAAGAGAGGCGCCAAAGAGAGAGAGAGCAAGTT 241
QY 333 TyrLeuGlyGlyTyrAspMetGluGluLysAlaAlaArgAlaTyrAspLeuAlaLeu 352
DB 242 TATCTAAGGGGTTATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
QY 353 LysTyrTyrGlyProSerThrHisIleAsnPheSerIleGluAsnTyrGlnValGlnLeu 372
DB 302 AGTATTTGGGGAGCCCTCTCTCATATTAATTTCCCATGTGAAATATCAAAACGAACTT 361
QY 373 GluGluMetIleAsnMetSerArgGlnGluLysValAlaHisLeuArgArgLysSerSer 392
DB 362 GAGGAAATTAACAACATGACGAGACAGAAATATGCTGCTATTACGAGAGAGAGAGAGAG 421
QY 393 GlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArg 412
DB 422 GGAATCTCAAGAGGGGCGCTCCATGTACAGAGAGATACAGGACCAACCAATGAGAGG 481
QY 413 ThrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheSer 432
DB 482 TGGCAAGCTCGAATGGAAGGGGTACCGGAAACCAAGATCTATATCTTGGAAACCTTTAGT 541
QY 433 ThrGlnGluGluAlaAlaGluAlaTyrAspValAlaAlaIleLysPheArgGlyAlaAsn 452
DB 542 ACCCAAGAGAGAGAGAGCTGACCTATAGCATGTGCTGATTAATTCAGAGAGATTAAT 601
QY 453 AlaValThrAsnPheAspIleSerArgTyrAspValGluArgIleMetAlaSerSerAsn 472
DB 602 GCTGTCACTAAGCTTGTATTAACAAGATATGACGTGGAAAAAATATATGAGAGCAATTAAC 661
QY 473 LeuLeuAlaGlyGluLeuAlaArgArgLysLysAspAsnAspProArgAsnLysAspIle 492
DB 662 CTTCTTAAGAGTGAACAGCTTAAGCGGAAAGAGAGATGAT----- 703
QY 493 AspTyrAsnLysSerValIleThrSerValAsn-----AsnGluGluThr 507
DB 704 GATGAAGACTAGAGAGAGGCTACA---GTTAACCAAAAACCTTCACTCAAGAAAC 760
QY 508 ValGlnValGlnAlaGlyAsnAsnAsnGlnAsnAspSerGluTyrIleMetValIleu 527
DB 761 GTTCTGATGCGAG-----AAAAGATGCAAGGCACTAGATGAGAGAGAGTTCAG 811
QY 528 PheAsnHisProSerGlnGlnGlnGlnAlaAsnGlyAsnGlySerAspGlnIleMet 547
DB 812 TTT-----CCATCCTCCCAACAGCTT-----GATCAGATATCCTCAAGAAATGAG 856
QY 548 AsnGlyGlyAsnTyrArgAsnSerAlaPheSerMetAlaLeu----- 561
DB 857 GGTGT-----AGAACTCAGCCCTTCTCAACAGCTTTAGATATAAGTITTCATAA 907
QY 561 ----- 561
DB 908 CAAGTTAGAGAGAGAGAGAACATGACAGACAGACATATGCTCATTTTGGAAACA 967

```

QY 562 -----GlnaspLeuIleGly 566  
 DB 968 AAGCAGCGGATTCTCAGAGGGCGCTCAGTACAAAGAGTAAACAAGGGCTCACAGGC 1027  
 QY 566 ----- 566  
 DB 1028 AAGCATTTACACATGACCACTTCATGGAAGGTCAAGCTGAGATGTGTAGTGAAA 1087  
 QY 567 -----IleaspSerValGlySerGly---GlnHis----- 575  
 DB 1088 CTAAACAAGATCTA-TATCTTGGCAGCGGTACTCAACATGAGAGAGCTGAAGCTATAG 1146  
 QY 575 ----- 575  
 DB 1147 AATTGCTGCTATTAAATTACAGAGGCGAATGCTGCTCAACTTGACATAACAGATA 1206  
 QY 575 ----- 575  
 DB 1207 TGACGTGTATGATGACGCAAGCAAAACCAACCAAGTAGAGCTGTTCAGCGTCC 1266  
 QY 575 ----- 575  
 DB 1267 CCGCCCCAGACAAATGATCTCAGTGAAGTGAAGTCAAGACTCAGCCCTTCTC 1326  
 QY 576 -----AsnMetLeuaspGlySerSerIleGlyThrHis 587  
 DB 1327 AAGGACTTAGATACATGTTTCTCAACCAAGTTGAGGAAAGGAGCAATGGGAAACACAC 1386  
 QY 588 PhaseIAsnThrSerSerLeuValThrSerLeuSerSerSerArgGluIleAspProGlu 607  
 DB 1387 TTGTCAAAATCCTTCTCTCTGCGTGAACAAGTTGAGTCAAGTCAAGAGAGAGCCAGAT 1446  
 QY 608 LysArgGlyProSerLeuLeuPheProMetProProMetGluThrIleValAsnPro 627  
 DB 1447 AAGACAAAGCATGCCAATGCTCTTGGATGATCCCTTCAACAGTGCACAAATATATG----- 1500  
 QY 628 IleGlyThrSerValThrSerTyrLeuProSerProThrValGlnMetArgProSerPro 647  
 DB 1501 -----GCTAACGCGGATCTCTGGATCTATCTTCCAACTCCAG-----ACT 1542  
 QY 648 AlalSerLeuSerHisLeuProValPheAlaSerThrProAsp 662  
 DB 1543 GCGCTTCTATGCTCAGATGCCAATTTTGTCTGCTGGACAGAT 1587  
 RESULT 15  
 US-10-059-911-32  
 ; Sequence 32, Application US/10059911  
 ; Publication No. US20030159180A1  
 GENERAL INFORMATION:  
 APPLICANT: Fischer, Robert L.  
 APPLICANT: Mizukami, Yukiko  
 TITLE OF INVENTION: The Regents of the University of California  
 FILE REFERENCE: 023070-090730US  
 CURRENT APPLICATION NUMBER: US/10/059,911  
 CURRENT FILING DATE: 2002-06-17  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 32  
 LENGTH: 1113  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: truncated or  
 OTHER INFORMATION: mutated ANT CDNA transgene ANTIDNC1  
 US-10-059-911-32  
 Alignment Scores:  
 Pred. No.: 3,57e-112 Length: 1113  
 Score: 1129.50 Matches: 245  
 Percent Similarity: 62.70% Conservative: 29  
 Best Local Similarity: 56.06% Mismatches: 58  
 Query Match: 32.25% Indels: 105

DB: 14 Gaps: 11  
 US-10-024-632-2 (1-663) x US-10-059-911-32 (1-1113)  
 QY 69 LeuThrValMetProLeuLysSerAspGlySerLeuCyS1LeuGluIleLeuLysArg 88  
 DB 1 ATGCTGTGATGGCAGCTCAGATCTGATGTTCTTCTTAATGGAAGCTTCAACAGA 60  
 QY 89 SerGlnThrGlnValMetValProThrSerSerProLysLeuGluAspPheLeuGlyGly 108  
 DB 61 TCTTCTCACTCGAATCAACCAATCAAGATTCTCTCCAAAGGTGAGATTTCTTGGG--- 117  
 QY 109 AlalThrMetGlyThrHisGlyTyrGlySerHisGluArgGly-----LeuSerLeuAsp 126  
 DB 118 -----ACCCATCAACAACACACAGTCAACAAGAACCCATGATCTTACCTTAGAT 168  
 QY 127 SerIleTyrTyrAsnSerGlnAsnAlaGluAlaGluProAsnArgAspLeuSerGln 146  
 DB 169 AGTTATTTCTACACACCACTCAT-----GAGCCCAACAGACTACAACTTTCMA 219  
 QY 147 ProPhe-----ArgGlnGlnGlyHisMetSerValGlnThrHisProTyrTyr 162  
 DB 220 GAGTTCTTAGCTTCCCTCAACACAGAACCAT----- 252  
 QY 163 SerGlyLeuAlaCySHisGlyLeuTyrGlnAlaProLeuGluGluThrThrHisGlu 182  
 DB 253 -----GAGGAGAACT-----AGA 267  
 QY 183 ThrHisValSerAspCySerSerLeuMetProGlnMetThrGluGlyLeuLysAsnTyr 202  
 DB 268 AATTACGGGATGAC-----CCTAGTTTGAACATGGA----- 300  
 QY 203 ValAlaProThrArgGluPheSerThrHisGlnValLeuGluGlnGlnMetAsnCyS 222  
 DB 301 -----GGGCTTTTAAATGA 315  
 QY 223 GlyMetGlyAsnGlyArgAsnGlyValSerLeuGlySerValGlyCyS1GluLeuGln 242  
 DB 316 GGGGATATGGGAAATTTCAA-----CAG 339  
 QY 243 SerLeuSerLeuSerMetSerProGlySerGlnSerSerCyS1ValThrAlaProSer--- 261  
 DB 340 TCAGTGAAGTATTCATGAGAGCCCTGCGTCAACATCTGACATGCTGCTCAGCCAC 399  
 QY 261 ----- 261  
 DB 400 CACCAACAAACCAAAACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 459  
 QY 262 -----GlyThrAspSerValAlaValAlaAspAlaLysLysArg 273  
 DB 460 CTGTGAGACAGAGCTTGGGTTGAGACGACGACCAATGCGGCTGCCAGAGAGAGG 519  
 QY 274 GlyHis-----AlaLysLeuGlyGlnGlyGlnProValHisArgLysSerIleAsp 290  
 DB 520 GACACAGAGAGATGTTGATGTTGTTGAGAAAGAGATGTTCTCAAGAAATCTATGAT 579  
 QY 291 ThrPheGlyGlnArgThrSerGlnTyrArgGlyValThrArgHisArgTyrThrGlyArg 310  
 DB 580 ACTTTTGGACAAAGCAACTTCTCAATACCGAGCGCTTACAAAGACATGATGAGTGTGAG 639  
 QY 311 TyrGluAlaHisLeuTyrAspAsnSerCyS1LysGlyGluGlnThrArgLysGlyArg 330  
 DB 640 TATGAGGTCATCTTGGACAAATGTTTCAAGAAAGAGGTACACAGTGAAGAGAGAGGA 699  
 QY 331 GlnValTyrLeuGlyGlyTyrAspMetGluGluLysAlaAlaArgAlaTyrAspLeuAla 350  
 DB 700 CAGATTATCTGGAGGTTATGATATGAGAGGAAAGCTGCTCAGACATATGATCTTCT 759  
 QY 351 AlalLeuLysTyrTyrGlyProSerThrHisIleAsnPheSerIleGluAsnTyrGlnAla 370  
 DB 760 GCACCTCAAGTACTGGGGTCCCTTACTCAACCAATTTCTTGGCGAGAAATATACGAAA 819  
 QY 371 GlnLeuGluGluMetLysAsnMetSerArgGlnGluTyrValAlaHisLeuArgArgLys 390

```
Db      820  GAGTTGAGACATGAGAACATGACTGACAGAAATATGTCACATTGAGAGGAG 879
Qy      391  SerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHis 410
Db      880  AGCAGTGGTTCTTCTTACGGGTCTTCCATCTATGAGAGCTCACAGACATCCAGCAT 939
Qy      411  GlyArgTpglnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThr 430
Db      940  GGAAGTGGCAGACGAGATTGTGAGAGTCGCTGAAACAAGATCTTACTTGGAAGT 999
Qy      431  PheSerThrGlnGlnGlnAlaAlaGlnAlaTyrThrAspValAlaAlaIleLysPheArgGly 450
Db      1000  TTGGAACCCAGAGAGAGCTGCAAGACTTACGATGTAGCAGCATTTAAGTCCGTTGC 1059
Qy      451  AlaAsnAlaValThrAsnPheAspIleSerArgTyrAspValGluArgIle 467
Db      1060  ACAATGTGTGACTTAATTGATATATCCAGAGTACGATGTGTATCGTATC 1110
```

Search completed: March 13, 2004, 11:23:41  
Job time : 483 secs



Wed Mar 17 08:18:56 2004

us-10-024-632-2.rst

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 13, 2004, 00:45:08 ; Search time 2976 Seconds

(without alignments)  
6652.764 Million cell updates/sec

Title: US-10-024-632-2  
Perfect score: 3502  
Sequence: 1 MKRINSNNNTDGNHNLG.....RSPALISHLPEFASWTD 663

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US10024632/runat\_09032004\_101208\_10776/app.query.fasta\_1.839  
-DB=EST -OMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCH=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blcsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=ptc -NORM=ext -HASPRTZ=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US10024632 @CGN 1.1 2.135 @runat\_09032004\_101208\_10776 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEUDERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_estnum:\*  
3: em\_estin:\*  
4: em\_estma:\*  
5: em\_estcov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gD\_estcl:\*  
10: gD\_est2:\*  
11: gD\_hcc:\*  
12: gD\_est3:\*  
13: gD\_est4:\*  
14: gD\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gD\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1103	31.5	752	10	AM348127
2	920	26.3	599	14	CA230632
3	911.5	26.0	1753	11	AY109146
4	910	26.0	558	14	CA232774
5	909	26.0	545	14	CD475882
6	895	25.6	769	14	CA783156
7	879	25.1	774	14	CF510900
8	855	24.4	882	14	CK267021
9	854	24.4	555	12	BJ188928
10	854	24.4	558	12	BJ178045
11	845	24.1	748	14	CF807326
12	840.5	24.0	590	12	BM094116
13	807	23.0	644	10	AM200688
14	792	22.6	500	10	AM780688
15	790.5	22.6	552	12	BU964897
16	787.5	22.5	593	13	BM307997
17	779	22.2	1160	14	CK206573
18	758.5	21.7	679	12	BJ312281
19	750.5	21.4	690	13	CA094356
20	750	21.4	720	13	BQ864461
21	746.5	21.3	631	13	CA103041
22	718	20.5	457	13	BU828154
23	715	20.4	591	14	CF244784
24	711.5	20.3	697	14	CA189006
25	708.5	20.3	752	14	CA238822
26	708.5	20.2	640	13	CA100375
27	708.5	20.2	674	10	BF647766
28	695.5	19.9	523	12	BF6790852
29	694.5	19.8	513	13	BQ122372
30	690.5	19.7	585	12	BM086088
31	682.5	19.5	606	13	BU825419
32	678	19.4	704	13	BQ625052
33	678	19.4	1897	11	AY103852
34	676	19.3	935	14	CK261980
35	650	18.6	907	14	CK289457
36	648	18.5	459	10	BE800260
37	647	18.5	728	14	CF035870
38	646	18.4	420	9	AJ475492
39	643.5	18.4	419	13	BU820600
40	643	18.4	694	14	CF036846
41	640.5	18.3	545	12	BI974354
42	640	18.3	641	14	CD879292
43	639.5	18.3	842	14	CF446573
44	638.5	18.2	737	14	CD832004
45	636	18.2	500	9	AI442942

# ALIGNMENTS

RESULT 1  
AM348127/c  
LOCUS  
DEFINITION  
AM348127  
SEQUENCE  
AM348127  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
752 bp mRNA linear EST 04-OCT-2000  
Gm210001A21A2 Gm-r1021 Glycine max cDNA clone Gm-r1021-4 3', mRNA  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 752)  
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,  
Expanding, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L., and Lewin, H.  
A Functional Genomics Program for Soybean (NSF 9872565)  
Unpublished (1999)  
Other Ests: A1444013  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l.vodkin@uiuc.edu  
This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134. For further information  
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or info@genome  
systems.com web site: www.genomesystems.com  
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.

## FEATURES

Location/Qualifiers  
1..752  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultiivar="Williams"  
/db\_xref="taxon:3847"  
/clone="Gm-r1021-4"  
/isue\_type="root"  
/lab\_host="XL10-Gold"  
/clone\_id="Gm-r1021"  
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:  
XhoI; Library Gm-r1021 is a sequence-driven, retracted set  
of the original library Gm-cl004 which was prepared from  
root cDNA. The mRNA was isolated from entire roots of 8  
day old 'Williams' seedlings which were propagated on  
paper towels with distilled water. Stragene's cDNA  
Synthesis Kit (catalog #200401) was used to synthesize  
the cDNA. The Gm-cl004 library was constructed by Dr.  
Paul Keim & Virginia H. Coryell, Department of Biology,  
Box 6640, Northern Arizona University, Flagstaff, AZ  
86011, email: Paul.Keim@uau.edu, Virginia.Coryell@uau.edu.  
The contig analysis to select unique genes was performed  
by the laboratory of Ernest Retzel, Computational Biology  
Centers, University of Minnesota,  
http://www.cb.cnm.edu/ResearchProjects/Soybean/index.html  
Retracking was performed by Genome Systems, St. Louis,  
http://www.genomesystems.com, and sequencing by the Keck  
Center for Comparative and Functional Genomics,  
University of Illinois,  
http://www.life.uiuc.edu/biotech/keck.html."

## ORIGIN

## Alignment Scores:

Pred. No.: 3.58e-97 Length: 752  
Score: 1103.00 Matches: 214  
Percent Similarity: 95.11% Conservative: 0  
Best Local Similarity: 95.11% Mismatches: 11  
Query Match: 31.50% Indels: 0  
Gaps: 0

US-10-024-632-2 (1-663) x AM348127 (1-752)

QY 439 GUAUATYRAspValAlaAlaIleLysPheArgGlyAlaAsnAlaValThrAspPheAsp 458  
DB 750 GAAGCATACGATGATGANNCCGATCAATTTCCGCGGCGNNNGCAGTACAACTTTGAC 691  
QY 459 ILeSerArgTYRAspValGluArgIleMetAlaSerSerAsnLeuValGlyGluLeu 478  
DB 690 ATTTCANNATACGANTGTNNNNNGAATCATNNNNAGTACCAATCTCTCGNNGGGGAGCTT 631  
QY 479 ALaArgArgLysAlaSerAspAspProArgAsnLysAspIleAspTYRAsnLysSerVal 498

DB 630 GCAAGCGCTAAGAAAGATACGATCTAGAAACAAGACATACACTCAACAAGAGTGA 571  
QY 499 ValThrSerValAsnArgGluThrValGlnValGlnIleGlyAsnAsnAsnArgGlu 528  
DB 570 GTAACAACTGTGAACAAATGAGAAACGGTTCACCTTCAAGCAGGAAACAAATATAGAA 511  
QY 519 AsnAspSerGluThrLysMetValLeuPheAsnHisPProSerGlnGlnGlnAlaAsn 538  
DB 510 AACGACTCAGAGTGAGATGATGTTTATTATTACACCCCTTCACAGCAGCAACAGCAAT 451  
QY 539 GlyAsnGlySerAspGlnLysIleMetAsnCyseGlySerValArgAsnSerAlaPheSer 558  
DB 450 GGCAATGGCACTGACCAAAAAATATGACCTGTGAAATTACAGAAACAGTGCATTTTCT 391  
QY 559 MetAlaLeuGlnAspLeuIleGlyIleAspSerValGlySerGlyGlnHisAsnMetLeu 578  
DB 390 ATGGCCCTACAAAGATCTATTGGATTGATTCGGGGGTTCTGGGACAGCATATATGCTG 331  
QY 579 AspGluSerSerLysIleGlyThrHisPheSerThrThrSerSerLeuValThrSerLeu 598  
DB 330 GACGAGTCTACCAAAATGGACATCTTTTCAACACAGTCATGCTGGTGAACAAGTTTA 271  
QY 599 SerSerSerArgGluAlaSerProGluLysArgGlyProSerLeuLeuPheProMetPro 618  
DB 270 AGCACTCAAGAGAGGCTAGTCTCTCGAAGAAAGGGTCCCTCGCTTCTCCATGCT 211  
QY 619 ProMetGluThrLysIleValAsnProIleGlyThrSerValThrSerTripleProSer 638  
DB 210 CCAATGAAACAACAAATGTGAACCCCATTTGTACACATGTTACTCTTGGCTACCCCTCA 151  
QY 639 ProThrValGlnMetArgProSerProAlaIleSerLeuSerHisLeuProValPheAla 658  
DB 150 CCANNGGTCAATAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 91  
QY 659 SerTPThrAspThr 663  
DB 90 TCTTGAGCTGATACT 76

RESULT 2  
CA230632 599 bp mRNA linear EST 25-SEP-2003  
LOCUS SCUFLIC06D04.9 Saccharum officinarum f33 Saccharum officinarum  
DEFINITION cDNA clone SCUFLIF3C06D04 5', mRNA sequence.  
ACCESSION CA230632  
VERSION CA230632.1 GI:35294356  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Bukayeva, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoidae; Andropogoneae; Saccharum.  
REFERENCE 1 (bases 1 to 599)  
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
COMMENT Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: pattina@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: C06 row: D column: 04  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1..599  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"

/clone="SCJFFL3C06D04"  
/lab host="DH108"  
/note="Organ: Base of developing inflorescence (5cm-long);  
vector: pSport1; Site\_1: SalI; Site\_2: NotI; An  
undirectional cDNA library generated from [base of  
developing inflorescence (5cm-long)]. cDNA was prepared  
from polyA+ mRNA using Superscript Plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a Sephadose Cl-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
<http://suceet.lad.ic.uticamp.br/public>

## ORIGIN

Alignment Scores:  
Pred. No.: 2,08e-79 Length: 599  
Score: 920.00 Matches: 168  
Percent Similarity: 97.33% Conservative: 14  
Best Local Similarity: 89.84% Mismatches: 5  
Query Match: 26, 27% Indels: 0  
DB: 14 Gaps: 0

US-10-024-632-2 (1-663) x CA230632 (1-599)

QY 301 GYVAThrrArgHisArgTTPThrGlyArgTyrGluAlaHisLeuTTPAspAsnSerCys 320  
DB 2 GGCCTCACAGGAGCTAGTGGACAGGAGGATAGAGCGCGACCTGTGGACAAACAGCTCC 61  
QY 321 LysLysGluGluGlnThrArgLysGlyArgGlnValTyrLeuGlyTyrAspMetGlu 340  
DB 62 AAGAGGAGAGGCGCACACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121  
QY 341 GlnLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTTPGlyProSerThrHis 360  
DB 122 GAGAGAGCGGAGAGGAGCTTACAGCTCGCGCGCTCAAGTACGAGGCGCTTCCACCCAC 181  
QY 361 LLeasnHeserLeuGluAsnTyrGlnValGlnLeuGlnMetLysAsnMetSerArg 380  
DB 182 ATCAACTTCCCGCTGAGAGACTACAGAGAGGAGTGAAGAGAGAGAGAGAGAGAGAG 241  
QY 381 GlnGlnTyrValAlaHisLeuArgLysSerSerSerGlyPheSerArgGlyAlaSerile 400  
DB 242 CAGAGAGTACGTGGCTCACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301  
QY 401 TyrArgLysValThrArgHisLeuGlnHisGlyArgTTPGlnAlaArgTLeuArgVal 420  
DB 302 TACCGAGAGAGTCAAG 361  
QY 421 AlaGlyAsnLysAspLeuTyrLeuGlyTyrPheSerThrGlnGlnGlnAlaAlaGlnAla 440  
DB 362 TCCGAG 421  
QY 441 TyrAspValAlaAlaAlaLeuLysPheArgGlyAlaAlaAlaValThrAsnPheAspLiser 460  
DB 422 TACAGCATCGCGCGATCAAGTTCGGGGGCGCTCAACCGCTCAACCTTCGATCAGCAG 481  
QY 461 ArgTyrAspValGluArgLysLeuMetAlaSerSerAsnLeuLysGlnGlnAlaArg 480  
DB 482 CGCTACGAGCGTCAAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541  
QY 481 ArgLysLysAspAsnAspPro 487  
DB 542 CGCAG 562

RESULT 3  
AT109146  
LOCUS AY109146 1753 bp mRNA linear HTC 17-OCT-2002  
DEFINITION Zea mays PCO137288 mRNA sequence.  
ACCESSION AY109146  
VERSION AY109146.1 GI:21212582  
KEYWORDS HTC.

SOURCE  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS  
Hainey, C.F., Dolan, M., Mao, G.H., Vogel, J.M., Whitsitt, M.S.,  
Arthur, L.W., Hainey, W., Morgante, M., and Tingey, S.V.  
Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 1753)  
Coe, E.H.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSU, [maizegenetics.org](http://maizegenetics.org); ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

FEATURES  
source  
1..1753  
location/Qualifiers  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/clone\_id="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

## ORIGIN

Alignment Scores:  
Pred. No.: 9.4e-78 Length: 1753  
Score: 911.50 Matches: 205  
Percent Similarity: 94.85% Conservative: 44  
Best Local Similarity: 45.15% Mismatches: 88  
Query Match: 26, 03% Indels: 117  
DB: 11 Gaps: 11

US-10-024-632-2 (1-663) x AY109146 (1-1753)

QY 187 AppCysSerSerLeuMetProGlnMetThrGlnGlyLeuLysAsnTTPValAlaProThr 206  
DB 42 GACATGAGCTCAGCTTATCCCAAC-----CATGGCTCTCTCTCC 83  
QY 207 ArgGlnPheSerThrHisGlnGlnValLeuGlnGlnMetAsnCysGly----- 223  
DB 84 CTCTCCACACATACCAACAGAGCTTCTCAAGCTTCTTAACTCTCCGTAATCCT 143  
QY 224 MetGlyAsnGluArgAsnGlyValSer-----LeuGly 234  
DB 144 CTGGAG 203  
QY 235 SerValGly----- 237  
DB 204 GGCCTGCTGGTGGCTGGCGCCCGCGAGAGCGGCGGTGAGATCAGATCAGATCAGAGCTT 263  
QY 238 ---CysGlyGlnLeuGlnSerLeuSerLeuSerMet----- 248  
DB 264 GTGTGGGAGAGCTGGAGAGATCACAGCAGAGCTTCTTGGCCACTACCCGCGGCGCA 323  
QY 249 -----SerProGlySerGlnSerSerCysValThrAlaProSerGly 262  
DB 324 GCTGGAG 377

















## ORIGIN

Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

## Alignment Scores:

Pred. No.:	1 266-71	Length:	590
Score:	840.50	Matches:	156
Percent Similarity:	90.77%	Conservative:	21
Best Local Similarity:	80.00%	Mismatches:	17
Query Match:	24.00%	Indels:	1
DB:	12	Gaps:	1

US-10-024-632-2 (1-663) x BW094116 (1-590)

```

QY 291 ThrPheGlyGlnATGThrSerGlnTyrArgGlyValThrArgHisArgTPTThrGlyArg 310
DB 1 ACATTCCGGGCGAGAAACATCCATATATCGGAGTAAACGGACATATAGTACTGGAAAG 60
QY 311 TyTGlnAHisLeuTTPASPASerCysLeuTyrGlnGlnThrArgGlyArg 330
DB 61 TATTAAGCTACCTTTGGGATATAGCTGTAGAAAGGAGGCAATCAAGAAAGACGC 120
QY 331 GlnValTyrLeuGlyGlyTyrAspMetGlnGlnValAlaAlaArgAlaTyrAspLeuAla 350
DB 121 CAAATTATTTGGGTGATATGATTAAGAAAGAAAGCACTAGTCTTATGATTAGCT 180
QY 351 AlaLeuTyrTyrTPGlyProSerThrHisLeuAspSerLeuGlnValTyrGlnVal 370
DB 181 GCACCTGAAGTACTGGGAGCATCCACCACTACCACTTCCATTACTATGAGAG 240
QY 371 GlnLeuGlnGlnMetLeuAspMetSerArgGlnGlnTyrValAlaHisLeuArgArg 390
DB 241 GAATTGATGAATGAAGAAACATGACCGGACAAATTTGTCTGCTCATTAGAGAGAA 300
QY 391 SerSerGlyPheSerArgGlyAlaSerLeuTyrArgGlyValThrArgHisLeuGlnHis 410
DB 301 AGCAGTGGTTCTCCAGGGGTCATCAATGATGTGAGTACCAAGGCATCCCAAC 360
QY 411 GlyArgTTPGlnAlaArgGlyValArgValAlaGlyAsnTyrAspLeuTyrLeuGlyThr 430
DB 361 GGAGATGGCAAGCAAGAAATTTGGCAGTTCCAGAAACAAAGATCTTTCTGGAGACT 420
QY 431 PheSerThrGlnGlnGlnValAlaGlnValTyrAspValAlaAlaAlaLeuPheArgGly 450
DB 421 TTCACTACTGAAGAGAGAGCTGCTGAGACATACATAGCTGCGATTAAGTTCAAGAGT 480
QY 451 AlaAlaAlaValAlaThrAspPheAspLeuSerArgTyrAspValGlnArgGlnMetAlaSer 470
DB 481 CTCACGCTGTCCAAACCTTGACATGAGCCGCTACGACGTGAAGCAACCATCTTGAAGC 540
QY 471 SerAsnLeu---LeuAlaGlyGlnLeuAlaArgArgTyrAsp 484
DB 541 AACACTCTCCCAATGAGAGAGAGAGCTGCAAAAGCTGCAAGAAAGAA 585

```

RESULT 13  
 AM200688  
 LOCUS  
 DEFINITION  
 se92c07.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl027-229 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. 1'  
 mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 1 (bases 1 to 644)  
 Shomaker, R., Keim, P., Vodkin, L., Eipel, J., Corryell, V.,  
 Khanna, A., Bolla, B., Matra, M., Hillier, L., Kueba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steple, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

TITLE  
JOURNAL  
COMMENT

Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R., and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cu@resgen.com  
 Insert Length: 968 Std Error: 0.00  
 High quality sequence stop: 418.

FEATURES  
source

1..644  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl027-229"  
 /csize\_type="corylons of 3- and 7-day-old Williams  
 seedlings"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl027"  
 /note="Vector: Bluescript II SK+ Site 1: EcoRI; Site 2:  
 XhoI. This cDNA library was constructed from mRNA isolated  
 from corylons of 3- and 7-day-old Williams seedlings  
 which were propagated on paper towels with distilled  
 water. The corylons were flash-frozen in liquid  
 nitrogen, then lyophilized for 72 hours. Unequal amounts  
 of mRNA was used for cDNA synthesis. Stratagene's cDNA  
 Synthesis Kit (catalog number 200401) was used to  
 synthesize the cDNA. First-strand synthesis was  
 performed with 5-methyl dCTP, hence the ligated cDNA was  
 hemimethylated. A modification of Stratagene's  
 first-strand synthesis primer was used. An anchor  
 nucleotide (V=A, C, or G) was added to the 3' end of the  
 primer (GAGAGAGAGAGAGAGAGAGAGTGTGAGT)18 to anchor  
 the primer at the 5' end of the poly(A) tract. After  
 second-strand synthesis, the cDNA ends were filled in  
 with cloned phi DNA, ligated to EcoRI adaptors and  
 subsequently phosphorylated. The XhoI site within the  
 first-strand synthesis primer was then restricted by the  
 digestion with XhoI; all XhoI sites in the cDNA would be  
 protected by their hemimethylated status. The cDNA  
 constructs were size-fractionated with a 500 bp cutoff,  
 using Glycerol Life Technologies' cDNA Size Fractionation  
 column. The column eluent was then ligated into  
 Stratagene's pBluescript(tm) II XR predigested vector  
 (pBluescript II SK+) that has been digested with EcoRI  
 and XhoI, and phosphorylated by Stratagene. 97% of the  
 white and blue colonies appear to contain recombinant  
 plasmids with cDNA inserts, based on size (n=30). This  
 library was constructed by Dr. Paul Keim and Dr. Virginia  
 Corryell."

## ORIGIN

## Alignment Scores:

Pred. No.:	2 82e-68	Length:	644
Score:	807.00	Matches:	148
Percent Similarity:	92.61%	Conservative:	15
Best Local Similarity:	84.09%	Mismatches:	13
Query Match:	23.04%	Indels:	0
DB:	10	Gaps:	0

US-10-024-632-2 (1-663) x AM200688 (1-644)

```

QY 281 TyGlnProValHisArgTyrSerLeuAspThrPheGlyGlnArgThrSerGlnTyrArg 300
DB 114 CAGAAACCTTCTCCAAAGAAACCGTCGACACTTGGCCAAAGCACTCCATCTACGCC 173

```

QY 301 GlyValThrArgHisArgTyrThrGlyValArgTyrGluAlaHisLeuThrPaspAsnSerCys 320  
 Db 174 GGCGTCACCCCGACATGATGACGCGGAAGATACGAAGCTCATCTATGGACAAATAGTTGT 233  
 QY 331 LysIleValGluGlyGlnThrArgGlyArgGlnValTyrLeuGlyGlyTyrAspMetGlu 340  
 Db 224 AGAAGAAGAGCCCAAGCAGGAAGGAAGACAAAGTTTACCGGGTGGTTATGACAAGAA 293  
 QY 341 GluIleValAlaArgAlaTyrAspLeuAlaAlaLeuIleTyrTrpGlyProSerThrHis 360  
 Db 294 GATAAGCGACGACGAGCTTACCATCTCCAGCTCCAGTACTGGGGGCCCACTACACACC 353  
 QY 351 IleAsnProSerIleGluAsnTyrGlnValGlnLeuGluGluMetCysAspMetSerArg 380  
 Db 354 ACCAATCTCCATTCCTCACTCATATGAGAAGAACTGGAGGAGATGAAACATACCAAG 413  
 QY 381 GlnGluTyrValAlaHisLeuArgArgIleSerSerGlyPheSerArgIleValSerIle 400  
 Db 414 CAAGAGTTGTGTTCTCTCTACGAAGAAGACAGTGTCTCTAGGGGGCCCTCTATA 473  
 QY 401 TTTAAGGlyValThrArgHisArgHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgVal 420  
 Db 474 TACACAGAGTACACAGACACACACATGCGATGCGACGAGCATGAGCAAGATT 533  
 QY 421 AlaGlyAsnIleAspLeuTyrLeuGlyTyrPheSerThrGlnGluAlaAlaGluAla 440  
 Db 534 GCGGAACAAAGACCTCTACTTGAACTTTGACGACCAAGAAAGAGCTGTGAGGCC 593  
 QY 441 TTTAATPValAlaAlaIleLeuPheArgGlyAlaAlaAlaValThrAsn 456  
 Db 594 TATGACATGCTGCTATCAAAATTCAGCGCATTTAAATGACATGACACAAAC 641

RESULT 14 500 bp mRNA linear EST 03-DEC-2001  
 AM780688  
 LOCUS 517507.Y1 Gm-c1027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-c1027-7165 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. 7,  
 mRNA sequence.

ACCESSION AM780688  
 VERSION AM780688.1 GI:7795291  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max

REFERENCE 1 (bases 1 to 500)  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marras, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, W., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, W., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McGinn, R., Waterson, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp, 2130  
 South Memorial Parkway Huntsville, AL 35891 For further information  
 call: (800)-533-4363 or contact via email: ccs@resgen.com  
 High quality sequence stop: 412.

## FEATURES

source  
 1..500  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3947"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1027-7165"

## ORIGIN

Alignment Scores:  
 Pred. No.: 5,336-67 Length: 500  
 Score: 792.00 Matches: 145  
 Percent Similarity: 94.58% Conservative: 12  
 Best Local Similarity: 87.35% Mismatches: 9  
 Query Match: 22,62% Indels: 0  
 DB: 10 Gaps: 0

US-10-024-632-2 (1-663) x AM780688 (1-500)  
 QY 294 GlnArgThrSerGlnTyrArgGlyValThrArgHisArgTyrThrGlyValArgTyrGluAla 313  
 Db 3 GAGCGACACCTCCATCTACCGCGCGCTCACCCGACATAGATGACGCGGAAGATACGAAGCT 62  
 QY 314 HisLeuThrPaspAsnSerCysIleValGluGlnThrArgGlyArgGlnValTyr 333  
 Db 63 CATCTATGGACAAATGTTGTATGAGAAGGCGCAAGGAGGAAGGAAGCAAGTTTAC 122  
 QY 334 LeuGlyGlyTyrAspMetGluGluValAlaAlaArgAlaTyrAspLeuAlaAlaLeuIle 353  
 Db 123 CTGGGTGGTTATGACAAGAAGATAGGACGACGAGCTTACGATCTCCAGCTCTCAAG 182  
 QY 354 TyrTrpGlyProSerThrHisIleAsnProSerIleGluAsnTyrGlnValGlnLeuGlu 373  
 Db 183 TACTGGGCTCACTACACACCAACCACTTCCCATTCCTCACTATGAGAAGAACTGGAG 242  
 QY 374 GluMetCysMetSerArgGlnGluTyrValAlaHisLeuArgArgIleSerSerGly 393  
 Db 243 GAGATGAAGACATGACAGGACAAAGTTGTTGCTCTCTACGAAGAAGACAGTGGT 302  
 QY 394 PheSerArgIleValSerIleTyrArgGlyValThrArgHisArgHisGlyArgTyr 413  
 Db 303 TTCTCTAGGGGGCCCTCTATATACAGAGAGTACACACACACGACATGCGCATGG 362  
 QY 414 GlnAlaArgIleGlyArgValAlaGlyAsnIleAspLeuTyrLeuGlyTyrPheSerThr 433

/tissue\_type="cotyledons of 3- and 7-day-old Williams  
 seedlings"  
 /lab\_host="DH10B"  
 /clone\_id="Gm-c1027"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI. This cDNA library was constructed from mRNA isolated  
 from cotyledons of 3- and 7-day-old Williams seedlings  
 which were propagated on paper towels with distilled  
 water. The cotyledons were flash-frozen in liquid  
 nitrogen, then lyophilized for 72 hours. Unequal amounts  
 of mRNA was used for cDNA synthesis. Stratagene's cDNA  
 Synthesis Kit (catalog number 200401) was used to  
 synthesize the cDNA. First-stranded synthesis was  
 performed with 5-methyl dCTP, hence the ligated cDNA was  
 hemimethylated. A modification of Stratagene's  
 first-strand synthesis primer was used. An anchor  
 nucleotide (V-A, C, or G) was added to the 3' end of the  
 primer (GAGAGAGAGAGAGAGAGAACTGTTCAGG(T)18) to anchor  
 the primer at the 5' end of the poly(A) tract. After  
 second-strand synthesis, the cDNA ends were filled in  
 with cloned Pfu DNA, ligated to EcoRI adapters and  
 subsequently phosphorylated. The XhoI site within the  
 first-strand synthesis primer was then restricted by  
 digestion with XhoI; all XhoI sites in the cDNA would be  
 protected by their hemimethylated status. The cDNA  
 constructs were size-fractionated with a 500 bp cutoff,  
 using GIBCOBRL Life Technologies' cDNA Size Fractionation  
 column. The column eluent was then ligated into  
 Stratagene's pBluescript(II) M13 digested vector  
 (pBluescript II SK(+)) that has been digested with EcoRI  
 and XhoI, and phosphorylated by Stratagene). 97% of the  
 white and blue colonies appear to contain recombinant  
 plasmids with cDNA inserts, based on size (n=30). This  
 library was constructed by Dr. Paul Keim and Dr. Virginia  
 Coryell."

Wed Mar 17 08:18:56 2004.

us-10-024-632-2.rst

Page 13

[illegible]

Alignment Scores:		8.91e-67	Length:	552
Pred. No.:	790.50	Matches:	146	
Score:	91.21%	Conservative:	20	
Percent Similarity:	80.22%	Mismatches:	15	
Best Local Similarity:	22.57%	Indels:	1	
Query Match:	13	Gaps:	1	
DB:				

US-10-024-632-2 (1-663) x BU964897 (1-552)

QY	301	GLIValThArGHiAaGTTPThClYArGTyTGlUAlAhISLEuTPRaPaasenserQYs	320
DB	3	GGAGTAAACCGACATAGATGACCTGGAAAGATAGTAAGTCACTTTGGGATTAATAGCTGt	62
QY	321	LYALeSLUGlYGLThThArGhAsGlyArGInAlTyRLeuGlyGlyTYRAsPmeGlu	340
DB	63	AGAAAGGAAAGGCGATCAAGAAAGAGCCAAAGTTATTTGGGTGATGTATAAGAA	122
QY	341	GLULySAIAAlaArGAlaTyRAsPLeuAlaAlaLeuLySTyTTPGlyProSerThRHis	360
DB	123	GAATAAGCAGCTAGCTTATGATTATTAAGTCACTGAAGTACTGGGAGCATCCACACT	182
QY	361	ILAsnPheSerIleGluAsnTyRGlInAlGlnLeuGluGluMetLysAsnMetSerArg	380
DB	183	ACCAACTTTCCATTGTACTATAGAAAGAAATGGATGAAATGAACACATGACGCGA	242
QY	381	GLInGlyTyRAlaAlaHisLeuArGArGlyYsserSerGlyPheSerArGlyAlaSerIle	400
DB	243	CAGAAATTTTCTGCTCCATTGAAGAGAAAGCAAGTGTCTTCAGGGGTGATCAATG	302
QY	401	TYArGGLYAlThArGHisHisGlnHisGlyArGTTPGlnAlaArGIlleGlyArGVal	420
DB	303	TATCTGTGAGTTCAAGCAGCATCACCAACACGAGAGATGCAAGAAATTTGGCAGATT	362
QY	421	AlAGLYAsnLYAsPLeuTyRLeuGlyThRPheserThRGlInGluAlaAlaGluAla	440
DB	363	GCGAGAAACCAAAATTTTACTTTGGAACTTTCACTAGTGAAGAAAGGCGTGAAGCA	422
QY	441	TYRAsPValAlaAlaIleLYsPheArGGLyAlaAlaAlaAlaAlaThRAsnPheAspIleSer	460
DB	423	TACGACCTACTGCTCGATTAAGTTCAAGAGTCTCAACGCTGCACAAACTTTGATATAGC	482
QY	461	ArGTyRAsPValGluArGIlleWecAlaSerSerAsnLeu--LeuAlaGlyGluLeuAla	479
DB	483	CGCTAGACAGCTGAAGAACCATCTTGAAAGCAACAACCTCCCAATAGAGAGGAGAGCTGCA	542
QY	480	ArGArG 481	
DB	543	AAAGGT 548	

Search completed: March 13, 2004, 07:59:35  
 Job time : 2990 secs

```
Search completed: March 13, 2004, 07:59:35
Job time : 2990 secs
```

## ORIGIN

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 10:47:09 ; Search time 38 Seconds  
(without alignments)

3684.069 Million cell updates/sec

Title: US-10-024-632-2

Perfect score: 3502  
Sequence: 1 MKRINESNNDDGNNHNLG.....RSPSAISLHLPVAFASWTDT 663

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCIT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCITUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3502	100.0	663	US-10-024-632-2	Sequence 2, Appl
2	1715	48.0	665	US-10-024-632-4	Sequence 4, Appl
3	1408.5	40.2	642	US-10-024-632-6	Sequence 6, Appl
4	1349.5	38.5	669	US-10-024-632-9	Sequence 9, Appl
5	1338	37.9	555	US-10-059-911-2	Sequence 2, Appl
6	1319	37.7	555	US-10-059-911-21	Sequence 21, Appl
7	1313	37.5	555	US-10-059-911-22	Sequence 22, Appl
8	1307	37.3	548	US-10-059-911-5	Sequence 5, Appl
9	1233.5	35.2	452	US-10-059-911-13	Sequence 13, Appl
10	1224	35.0	473	US-10-059-911-9	Sequence 9, Appl
11	1180.5	33.7	585	US-10-024-632-11	Sequence 11, Appl
12	1129.5	32.3	370	US-10-059-911-15	Sequence 15, Appl
13	1042	29.8	308	US-10-059-911-10	Sequence 10, Appl
14	955.5	27.3	277	US-10-059-911-11	Sequence 11, Appl
15	948	27.1	574	US-10-374-780A-334	Sequence 334, App

16	947.5	27.1	205	US-10-059-911-17	Sequence 17, Appl
17	939.5	26.8	275	US-10-059-911-25	Sequence 25, Appl
18	939.5	26.8	415	US-10-374-780A-400	Sequence 400, App
19	927	26.5	558	US-10-374-780A-2496	Sequence 2496, App
20	891.5	25.5	498	US-10-374-780A-346	Sequence 346, App
21	889	25.4	314	US-10-059-911-23	Sequence 23, Appl
22	861	24.6	174	US-10-059-911-19	Sequence 19, Appl
23	762.5	21.8	357	US-10-059-911-14	Sequence 14, Appl
24	731.5	20.9	490	US-10-374-780A-1765	Sequence 1765, App
25	682	19.5	430	US-09-934-455-154	Sequence 154, App
26	682	19.5	430	US-10-286-264-104	Sequence 104, App
27	682	19.5	430	US-10-094-458A-6	Sequence 6, Appl
28	682	19.5	430	US-10-295-403-34	Sequence 34, Appl
29	682	19.5	430	US-10-374-780A-202	Sequence 202, App
30	681	19.4	430	US-10-094-458A-3	Sequence 3, Appl
31	675.5	19.3	443	US-10-374-780A-1311	Sequence 1311, App
32	666.5	19.0	313	US-10-225-067-140	Sequence 140, App
33	666.5	19.0	313	US-10-374-780A-2164	Sequence 2164, App
34	658.5	18.8	275	US-10-059-911-16	Sequence 16, Appl
35	638	18.2	415	US-10-374-780A-1314	Sequence 1314, App
36	633	18.1	399	US-10-374-780A-1313	Sequence 1313, App
37	614	17.5	622	US-10-374-780A-1743	Sequence 1743, App
38	611.5	17.5	255	US-10-024-632-13	Sequence 13, Appl
39	567	16.2	429	US-10-374-780A-1312	Sequence 1312, App
40	566	16.2	426	US-10-374-780A-1310	Sequence 1310, App
41	558.5	15.9	196	US-10-059-911-12	Sequence 12, Appl
42	532	15.2	264	US-10-059-911-24	Sequence 24, Appl
43	517	14.8	205	US-10-094-458A-8	Sequence 8, Appl
44	483.5	13.8	428	US-09-934-455-42	Sequence 42, Appl
45	483.5	13.8	428	US-10-374-780A-536-82	Sequence 82, Appl

#### ALIGNMENTS

RESULT 1					
US-10-024-632-2					
Sequence 2, Application US/10024632					
Publication No. US20020170093A1					
GENERAL INFORMATION:					
APPLICANT: Monsanto Technology LLC					
APPLICANT: He, Steve S.					
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION					
TITLE OF INVENTION: GROWTH AND USES THEREOF					
FILE REFERENCE: 38-21(51837)B					
CURRENT FILING DATE: 2001-12-19					
PRIOR APPLICATION NUMBER: US/10/024,632					
PRIOR FILING DATE: 2000-12-21					
NUMBER OF SEQ ID NOS: 33					
SOFTWARE: PatentIn version 3.1					
SEQ ID NO 2					
TYPE: PRT					
LENGTH: 663					
ORGANISM: Glycine max					
US-10-024-632-2					
Query Match					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MKRINESNNDDGNNHNLGFSLSPHMKMEATSAATVPTTFYMSPSQSHLSNFGMCYGVG	60		
DB	1	MKRINESNNDDGNNHNLGFSLSPHMKMEATSAATVPTTFYMSPSQSHLSNFGMCYGVG	60		
QY	61	ENGFNHSLITMPPKSPGSLCILEALRSOTQVWVPTSPKLBDFPGATGTHREYGSHE	120		
DB	61	ENGFNHSLITMPPKSPGSLCILEALRSOTQVWVPTSPKLBDFPGATGTHREYGSHE	120		
QY	121	RGSLDSITVYNSQNAEAQPNRDLISQPPROQGMVQTHFYVGLACHGLYQAPLEERT	180		
DB	121	RGSLDSITVYNSQNAEAQPNRDLISQPPROQGMVQTHFYVGLACHGLYQAPLEERT	180		

QY 181 KETHVSDSSIMPQMTBGLKNVVAIPTRFSTHQVLEQNMCKGNRNGVLSGVCCG 240  
DB 181 KETHVSDSSIMPQMTBGLKNVVAIPTRFSTHQVLEQNMCKGNRNGVLSGVCCG 240  
QY 241 LOSLSLMSPGSSCCTAPSGTDSVAADAKRGHAKLGQKQPHRKSIDTFGQRTSQYR 300  
DB 241 LOSLSLMSPGSSCCTAPSGTDSVAADAKRGHAKLGQKQPHRKSIDTFGQRTSQYR 300  
QY 301 GYTRHRTGRYEAHLMDNSCKKEGQTRKGRQVYIGYDMEKARAADLALKTWGPSTH 360  
DB 301 GYTRHRTGRYEAHLMDNSCKKEGQTRKGRQVYIGYDMEKARAADLALKTWGPSTH 360  
QY 361 INFSTENYQVLEEMKMSRQRYVAHLRKSISGFSRGASIRGVTRHHQGRWQARIGRY 420  
DB 361 INFSTENYQVLEEMKMSRQRYVAHLRKSISGFSRGASIRGVTRHHQGRWQARIGRY 420  
QY 421 AGNKDLYLGTSTQEEAAEAYDVAIKFRGANAVTNPDISRYDVERIMASSNLLAGELAR 480  
DB 421 AGNKDLYLGTSTQEEAAEAYDVAIKFRGANAVTNPDISRYDVERIMASSNLLAGELAR 480  
QY 481 RKDNDPRNKDIDYKSVYTSVNNETVQVQAGNNNNNDEEMQVLFNHSQQQOANGN 540  
DB 481 RKDNDPRNKDIDYKSVYTSVNNETVQVQAGNNNNNDEEMQVLFNHSQQQOANGN 540  
QY 541 GSDOKIMNGNYNSAFSMALQDLIGIDVSGQHMLDESSKIGTFHFSNTSLVTSISS 600  
DB 541 GSDOKIMNGNYNSAFSMALQDLIGIDVSGQHMLDESSKIGTFHFSNTSLVTSISS 600  
QY 601 SEASPEKRGESLIPMPMETKINPIGTSTSLPSPYQMRSPALISHLVPFASW 660  
DB 601 SEASPEKRGESLIPMPMETKINPIGTSTSLPSPYQMRSPALISHLVPFASW 660  
QY 661 TDT 663  
DB 661 TDT 663

## RESULT 2

US-10-024-632-4  
; Sequence 4, Application US/10024632  
; Publication No. US20020170093A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; APPLICANT: He, Steve S.  
; APPLICANT: Dotson, Stanton B.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
; FILE REFERENCE: 38-21(51837)B  
; CURRENT APPLICATION NUMBER: US/10/024,632  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,896  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-024-632-4

Query Match 49.0%; Score 1715; DB 13; Length 665;  
Best Local Similarity 54.4%; Pred. No. 2,2e-151;  
Matches 379; Conservative 84; Mismatches 166; Indels 68; Gaps 20;

QY 1 MKAINSNNTDDGNHNMIGFSLSPHM-----MKMTSAATVPTTFYMGPSQSHSN 52  
DB 1 MKMENDDNADLNQNNWIGFSLSPQMHNTGVSHSQPSAAEAVTSTSYHTTAP--LSS 58  
QY 53 GGNMNGV-GENGNFHSPLTVPPLKSDGSLCILEALKRSQTVVVPPTSPKLEDFLGATM 111  
DB 53 YGPFYIGLEANNVGLYSALPIMPPLKSDGSLVGLTILRSQQAQAWATSTPKLEDFLGATM 118

QY 112 GT-HEY-----GSHERGLSDSIYVNSQNAEAPNEDLLSQPERQ-----QGNMVOYTHPY 162  
DB 112 GT-HEY-----GSHERGLSDSIYVNSQNAEAPNEDLLSQPERQ-----QGNMVOYTHPY 162  
QY 119 GTPHHYEGSATETWPLSLDSVY--IQSRBDPNNQTVQNHVQHTSTNQOQOQOELQAY 177  
DB 119 GTPHHYEGSATETWPLSLDSVY--IQSRBDPNNQTVQNHVQHTSTNQOQOQOELQAY 177  
QY 163 SGLACHGLYQAPLEETTKEHVSQSSLMQ-----NTEGLXNVAPTRR--STHQ 213  
DB 163 SGLACHGLYQAPLEETTKEHVSQSSLMQ-----NTEGLXNVAPTRR--STHQ 213  
QY 178 STLRHNDML-----EGSRQSGTSDNNLHVQNMGGDDAVPVPLKSN--EVRNFQASHA 230  
DB 178 STLRHNDML-----EGSRQSGTSDNNLHVQNMGGDDAVPVPLKSN--EVRNFQASHA 230  
QY 214 QVLEQNMCKGNRNGV--SLGSVCGELQSLSLMSPGSSCCT-----APSGTDSV 266  
DB 214 QVLEQNMCKGNRNGV--SLGSVCGELQSLSLMSPGSSCCT-----APSGTDSV 266  
QY 231 H--EGMVIYPHEENAGSSGSLGMAYPDQLSLMSPGSSSVTSSHRASPAVDSV 288  
DB 231 H--EGMVIYPHEENAGSSGSLGMAYPDQLSLMSPGSSSVTSSHRASPAVDSV 288  
QY 267 AYDAKRGHAKLGQKQPHRKSIDTFGQRTSQRYGVTRHRTGRYEAHLMDNSCKKEGT 326  
DB 267 AYDAKRGHAKLGQKQPHRKSIDTFGQRTSQRYGVTRHRTGRYEAHLMDNSCKKEGT 326  
QY 289 AMDTKRGPEKVDQKQIYHRSIDTFGQRTSQRYGVTRHRTGRYEAHLMDNSCKKEGS 348  
DB 289 AMDTKRGPEKVDQKQIYHRSIDTFGQRTSQRYGVTRHRTGRYEAHLMDNSCKKEGS 348  
QY 327 RKGRQVYIGYDMEKARAADLALKTWGPSTHINFSTENYQVLEEMKMSRQRYVAH 386  
DB 327 RKGRQVYIGYDMEKARAADLALKTWGPSTHINFSTENYQVLEEMKMSRQRYVAH 386  
QY 349 RKGRQVYIGYDMEKARAADLALKTWGPSTHINFSTENYQVLEEMKMSRQRYVAH 408  
DB 349 RKGRQVYIGYDMEKARAADLALKTWGPSTHINFSTENYQVLEEMKMSRQRYVAH 408  
QY 387 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTQEEAAEAYDAI 446  
DB 387 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTQEEAAEAYDAI 446  
QY 409 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTQEEAAEAYDAI 468  
DB 409 LRKSSGFSRGASIRGVTRHHQGRWQARIGRVANCKDLYLGTSTQEEAAEAYDAI 468  
QY 447 KFRGANAVTNPDISRYDVERIMASSNLLAGELARCKNDPRNKDIDYKSVYTSV--NNE 505  
DB 447 KFRGANAVTNPDISRYDVERIMASSNLLAGELARCKNDPRNKDIDYKSVYTSV--NNE 505  
QY 469 KFRGANAVTNPDISRYDVERIMASSNLLAGELARCKNDPRNKDIDYKSVYTSV--NNE 528  
DB 469 KFRGANAVTNPDISRYDVERIMASSNLLAGELARCKNDPRNKDIDYKSVYTSV--NNE 528  
QY 506 ETVOQAGNNNNNDEEMQVLFNHSQQQOANGSDOKIMNGNYNSAFSMALQDLI 565  
DB 506 ETVOQAGNNNNNDEEMQVLFNHSQQQOANGSDOKIMNGNYNSAFSMALQDLI 565  
QY 529 EALMHQKCESEND--QMKVLY--QSSQLEQNPPTIE-----SDRTQSPAVLDMNF 580  
DB 529 EALMHQKCESEND--QMKVLY--QSSQLEQNPPTIE-----SDRTQSPAVLDMNF 580  
QY 566 GIDVSGQHMLDESSKIGTFHFSNTSLVTSISSREASPEKRGESLIPMPMETKIV 625  
DB 566 GIDVSGQHMLDESSKIGTFHFSNTSLVTSISSREASPEKRGESLIPMPMETKIV 625  
QY 581 -----HQVESSKARITHVSNPSLATSLSRBSRPSRTSLPMLSGMPTASKUL 631  
DB 581 -----HQVESSKARITHVSNPSLATSLSRBSRPSRTSLPMLSGMPTASKUL 631  
QY 626 NPIGTSVSWLPSPYQMRSPALISHLVPFASWTD 662  
DB 626 NPIGTSVSWLPSPYQMRSPALISHLVPFASWTD 662  
QY 632 ATNPNVNSWDPSPHLR-----PALTLPOMFVFAWTD 664  
DB 632 ATNPNVNSWDPSPHLR-----PALTLPOMFVFAWTD 664

## RESULT 3

US-10-024-632-6  
; Sequence 6, Application US/10024632  
; Publication No. US20020170093A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; APPLICANT: He, Steve S.  
; APPLICANT: Dotson, Stanton B.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
; FILE REFERENCE: 38-21(51837)B  
; CURRENT APPLICATION NUMBER: US/10/024,632  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: US 60/257,896  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-024-632-6

Query Match 40.2%; Score 1408.5; DB 13; Length 642;  
Best Local Similarity 46.5%; Pred. No. 1.1e-122;  
Matches 341; Conservative 64; Mismatches 149; Indels 179; Gaps 27;

QY 13 GNNHNMIGFSLSPMKMEATSAATVPTT-----F 41  
DB 13 GNNHNMIGFSLSPMKMEATSAATVPTT-----F 41  
QY 5 GGNMNGV-GENGNFHSPLTVPPLKSDGSLCILEALKRSQTVVVPPTSPKLEDFLGATM 111  
DB 5 GGNMNGV-GENGNFHSPLTVPPLKSDGSLCILEALKRSQTVVVPPTSPKLEDFLGATM 111  
QY 42 YMSPSQSHL-----SNFMCYGVGEN--GNFHSPLTVPPLKSDGSLCILEALKRSQTV 92  
DB 42 YMSPSQSHL-----SNFMCYGVGEN--GNFHSPLTVPPLKSDGSLCILEALKRSQTV 92

```

Db      65 LFPBPAQWVAPSPGYVYGAGYGDGTSTAGVYSHLPVMPFIKSDSLCIMEG----- 117
QY      93 VMPTSPKLEDELLGATWGTHTYGHSHERGLSDSYNSQNAEAPNDDLSQPPROG 152
Db      118 -MMSSSPKLEDELLGCGNSGSH-----DPATYISQGOEAD-----ASRAYOH 160
QY      153 HNSVQTPFYSGHAGHLYQAPLEETTKETHVSDSSIMPOMTEGLKMWVAPTRFEFSTH 212
Db      161 HOLV-----FVN-----YQPLTEAEMLQEAAPAPMEDAM-----AAAKNF----- 195
QY      213 QQVLEEQMGCMGNERNGVSLGVSQCEL-QSLSLSMSPGSO-SSCVT-APSGTSDVAD 269
Db      196 ---LVTSYGACYGNO-----EMPOPISLSMSPGSOSSSCVSAAPQOHOQWAV 240
QY      270 A-----KKRGAHLGQKQPVHRKSIDTFQORTSQYRGVTRHRM 307
Db      241 AAAAAADGGGNSNDGEGORVGKRGTKGQKQPVHRKSIDTFQORTSQYRGVTRHRM 300
QY      308 TGRYEALHMDNSCKEGQTKRGQVYLGVDMEKAPAYDLAALKYMGSTHINSTEIN 367
Db      301 TGRYEALHMDNSCKEGQTKRGQVYLGVDTEKAPAYDLAALKYMGSTHINSTEIN 360
QY      368 YQVLEEMKMSRQEVYVAHLRRKSSGFSRGASIRGVTTRHOGHQRQARIGRVAGNKDLY 427
Db      361 YRDEIEEMERTKQEVYVAHLRRSSGFSRGASIRGVTTRHOGHQRQARIGRVAGNKDLY 420
QY      428 LGTSTOEAAEAYDAVAIKFRGANAVTNFDSRYVERIMASSNLIGELARRK----- 483
Db      421 LGTSTOEAAEAYDAVAIKFRGLNAVTFDITRYVDKIMESSSLPGGAARKVAIEA 480
QY      484 --DNDPRNKIDYKSVTTSVNNETVQVQAGNNNNNDSEWKKVLENNHSSQOQA----- 537
Db      481 APDPVPIGREL-----GATEBASAAATVGT-----TWKRVL--HSSQOQAAC 523
QY      538 NGNSDQKIMNCGYRNASAFSMALQDIDISVSGGQNMLEDSSKI-GTHFSNTSLVT 596
Db      524 EATMDLQK-----GFMDDAS-ALHGIVGFVESAADEIDVPGKISGINFSSSLVT 577
QY      597 SLSSREASPEKRGPSLFFMPMETKIYVPIGTSVTS---WLPSPVQ---MRBPAP 649
Db      578 SLSSREASPEKRGPSLFFMPMETKIYVPIGTSVTS---WLPSPVQ---MRBPAP 630
QY      650 SLSHLPVFAWTD 662
Db      631 --AHLPVFAWTD 641

```

## RESULT 4

```

US-10-024-632-9
Sequence 9, Application US/10024632
Publication No. US20020170093A1
GENERAL INFORMATION:
APPLICANT: Monsanto Technology LLC
APPLICANT: He, Steve S.
APPLICANT: Dotson, Stanton B.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION A
TITLE OF INVENTION: GROWTH AND USES THEREOF
FILE REFERENCE: 38-21(51837)B
CURRENT FILING DATE: US/10/024,632
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/257,896
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 669
TYPE: PRT
ORGANISM: Oryza sativa
US-10-024-632-9

```

Query Match 38.5%; Score 1349.5; DB 13; Length 669;  
Best Local Similarity 45.3%; Pred. No. 46-117;

```

Matches 341; Conservative 66; Mismatches 142; Indels 203; Gaps 34;
QY      13 GNNHNLGFSLSPMKREAT-----SATVPTTFWSP 45
Db      18 GGVGAMLGFSLSPMATYACAGVDVGHNNHHVHOQOHHGGGLFYNPAVASFYGG 77
QY      46 SQ-----SLSNFMGCVGNGFNHSPULVMPKSDSLCTLEALK--RSQTVMPFSS 99
Db      78 GHDAVNTSAAGGSYTGAG-----FSSNPLSDSLCTLEALKRGDQEQGVVVSAS 129
QY      100 PKLEDFLG-GATWGTHTYGHSHERGLSD--SIYNSQNAEAPNDDLSQPPROGHN-- 154
Db      130 PKLEDFLGAGPAM-----ALSLDNSAFYGGCHG-----HHQHPAD 165
QY      155 --SVQTHFYSG--LACHGLYQAPLEETTKETHVSDSSIMPOMTEGLKMWVAPTRFEF 209
Db      166 GAVGAGDEHHGGGGFLQC-----AVIFGAGAG----- 192
QY      210 STHQVLEEQMNC-----GMGNERNGVSLGVSQ-----GELQSLISM- 248
Db      193 --HDAALVHDQSAANAAGMAAHGGGVDIANAAADV-CAAGPIIPFGHILHPLTSSMS 249
QY      249 SPGSGSCVT--APSGTSDVAVDA--KKRGA-KLGQKQPVHRKSIDTFQORTSQYRGV 302
Db      250 SAGSOSCVTVQAAAGAPYVAMDAVSKRGADARAGQKQPVHRKSIDTFQORTSQYRGV 309
QY      303 TRHWTGTYEALHMDNSCKEGQTKRGQVYLGVDMEKAPAYDLAALKYMGSTHIN 362
Db      310 TRHWTGTYEALHMDNSCKEGQTKRGQVYLGVDMEKAPAYDLAALKYMGSTHIN 369
QY      363 FSIENYQVLEEMKMSRQEVYVAHLRRKSSGFSRGASIRGVTTRHOGHQRQARIGRVAG 422
Db      370 FLEDYQSELEEMKMSRQEVYVAHLRRSSGFSRGASIRGVTTRHOGHQRQARIGRVAG 429
QY      423 NKDYLIGF-----STOEAAEAYDAVAIKFRGANAVTNFDSRYVERIM 468
Db      430 NKDYLIGFIFASAPAAPARAHAGTQEBAAEAYDAVAIKFRGLNAVTFDITRYVDKIL 489
QY      469 ASSNLIGELARRK-KDNDPRNKIDYKSVTTSVNNETVQVQAGNNNNNDSEWKKVLE 527
Db      490 ESSTLLPGELARRKKGVDGGG-----AAVADAALALVQAG-----NVLEMKAT 535
QY      528 -----FNHPSQOQANGNSQKIMNCGYRNASAFSMALQDIDISVSGGQNMLEDSS 581
Db      536 AALPAAARTEQOQHHGGGHO--HDDLPRDASV-LQDITVTDVDAAGAPR----- 585
QY      582 SKIGTHFSNTSLVTSLSSREASPEKRG-----SLFPMPMETKIYVPIGTSV 632
Db      586 ---APHMSMA---TSLGNSREGSPD-RGVGGGGGGVLTATIFAKPAASKLYSFV--PL 636
QY      633 TSWL-PSPTVQMRPSPA-TLSHLPVFAWTD 662
Db      637 NTWASPSPAVSSVPFARAGVISIAHLPFAWTD 668

```

## RESULT 5

```

US-10-059-911-2
Sequence 2, Application US/10059911
Publication No. US20030159180A1
GENERAL INFORMATION:
APPLICANT: Fischer, Robert L.
APPLICANT: Mizukami, Yukiko
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants
CURRENT FILING DATE: US/10/059,911
CURRENT FILING DATE: 2002-06-17
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 555
TYPE: PRT
ORGANISM: Arabidopsis thaliana

```



## FEATURE:

OTHER INFORMATION: Arabidopsis AINTEGUMENTA (ANT)  
US-10-059-911-2

Query Match 37.9%; Score 1328; DB 14; Length 555;  
Best Local Similarity 43.5%; Pred. No. 3,1e-115;  
Matches 314; Conservative 69; Mismatches 113; Indels 226; Gaps 22;

```

QY 1 MKRINENNTDDGNHNMWLGFSLSPH-MKM-----EATSAATVPTTFYM 43
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 1 MKSFCDDNDNNHNTNLLGFSLSNMWKGGRGREALYSSSTSSAATSSSVPPQLV 60
QY 44 SPSQSHLSNFGMCYGVGENGFHSPLYTMPLKSDGSLCILKRSQTVWPTSSPKLE 103
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 61 G---DNTSNFGVCYGSNPGGYSMSVMPLRSDGSLCMEALNRSSSHNHODSSPKYE 117
QY 104 DFLGATWGTGHEYSHERG--LSLDSIYNSQNAEAQPNRDLISQPF---RQGHMSVQ 157
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 118 DFFG---THNNTSHKEMDLSDLSLFYNTTH---EPNTTNGFEFFSPQTRNH--- 166
QY 158 THPYSGLAHGLYQAPLEETTKETHVSDCSLMPONTGKKNWVAPTRFSTHQVLE 217
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 167 ---EERT--RNYGND---PSLTHG----- 182
QY 218 QQNNCGMNERNGVSLGSGGEL-QSLSLSPSGSGSCVTAPS----- 261
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 183 ---GSFNVGV---YGFQOQSLSLSPSGSGSCITGSHHQQNQNHQSQNH 229
QY 262 ---GTDVAVDAKKRGH---AKLGOKQPVHRKSIDTGGRTSQRYGTR 304
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 230 QQISEALVETSVGFETTMAAKKKRQGEDVAVVQKQIVHRKSIDTGGRTSQRYGTR 289
QY 305 HRWTRIRYEAHLMDNSCKEKGQTRKRGQVYLGVDMEKARAAYDLALKYWGSPSTHIFS 364
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 290 HRWTRIRYEAHLMDNSCKEKGSRKRGQVYLGVDMEKARAAYDLALKYWGSPSTHIFS 349
QY 365 IENYQVQLEEMKMSQOEYVAHLRRKSSGFSRGASITRGVTRHHQGRQARIGRVAGNK 424
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 330 AENYQKEIEDMKMTRQOEYVAHLRRKSSGFSRGASITRGVTRHHQGRQARIGRVAGNK 409
QY 425 DLYLGTFTQOEBAEAAYVAALKFRGANAVTNFDISRVDYRIMASNTLAGELARKKD 484
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 410 DLYLGTFTQOEBAEAAYVAALKFRGTNAVNTFDITRYVDIMSNNTLSGELARRN-- 467
QY 465 NDPRKMDIDYNSVTVSNNEETVQVQAGNNNNENDSEMKVTLFNHPSCQQOANGNGSDQ 544
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 468 ---NNSIV--VRNTE-----DQ 479
QY 545 KIMNCGNRYNSAFSMALQDLIGIDVSGQHMLDESSKIGTHFSNTSSLYTSLSSREA 604
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 480 TALNA-----VEGGSNREV 494
QY 605 SPEKRGPSL--LFPMPMETKIV-NPIGNSVTSMLESPFVQMRPSPALISLHPVFAASMT 661
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 495 STPERLISFPAIPALPQVAVQKFGSGMNGMSPMTSNPAELK-TVALTLQMPVFAAMA 553
QY 662 DT 663
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 554 DS 555

```

## RESULT 6

US-10-059-911-21  
Sequence 21, Application US/10059911  
Publication No. US20030159180A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
FILE REFERENCE: 023070-090730US  
CURRENT APPLICATION NUMBER: US/10/059,911  
CURRENT FILING DATE: 2002-06-17

NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 21  
LENGTH: 555  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE: Description of Artificial Sequence: structurally  
OTHER INFORMATION: altered ANT protein AINTEGUMENTA  
US-10-059-911-21

Query Match 37.7%; Score 1319; DB 14; Length 555;  
Best Local Similarity 43.2%; Pred. No. 2.2e-114;  
Matches 312; Conservative 70; Mismatches 114; Indels 226; Gaps 22;

```

QY 1 MKRINENNTDDGNHNMWLGFSLSPH-MKM-----EATSAATVPTTFYM 43
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 1 MKSFCDDNDNNHNTNLLGFSLSNMWKGGRGREALYSSSTSSAATSSSVPPQLV 60
QY 44 SPSQSHLSNFGMCYGVGENGFHSPLYTMPLKSDGSLCILKRSQTVWPTSSPKLE 103
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 61 G---DNTSNFGVCYGSNPGGYSMSVMPLRSDGSLCMEALNRSSSHNHODSSPKYE 117
QY 104 DFLGATWGTGHEYSHERG--LSLDSIYNSQNAEAQPNRDLISQPF---RQGHMSVQ 157
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 118 DFFG---THNNTSHKEMDLSDLSLFYNTTH---EPNTTNGFEFFSPQTRNH--- 166
QY 158 THPYSGLAHGLYQAPLEETTKETHVSDCSLMPONTGKKNWVAPTRFSTHQVLE 217
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 167 ---EERT--RNYGND---PSLTHG----- 182
QY 218 QQNNCGMNERNGVSLGSGGEL-QSLSLSPSGSGSCVTAPS----- 261
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 183 ---GSFNVGV---YGFQOQSLSLSPSGSGSCITGSHHQQNQNHQSQNH 229
QY 262 ---GTDVAVDAKKRGH---AKLGOKQPVHRKSIDTGGRTSQRYGTR 304
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 230 QQISEALVETSVGFETTMAAASSTRQGEDVAVVQKQIVHRKSIDTGGRTSQRYGTR 289
QY 305 HRWTRIRYEAHLMDNSCKEKGQTRKRGQVYLGVDMEKARAAYDLALKYWGSPSTHIFS 364
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 290 HRWTRIRYEAHLMDNSCKEKGSRKRGQVYLGVDMEKARAAYDLALKYWGSPSTHIFS 349
QY 365 IENYQVQLEEMKMSQOEYVAHLRRKSSGFSRGASITRGVTRHHQGRQARIGRVAGNK 424
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 330 AENYQKEIEDMKMTRQOEYVAHLRRKSSGFSRGASITRGVTRHHQGRQARIGRVAGNK 409
QY 425 DLYLGTFTQOEBAEAAYVAALKFRGANAVTNFDISRVDYRIMASNTLAGELARKKD 484
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 410 DLYLGTFTQOEBAEAAYVAALKFRGTNAVNTFDITRYVDIMSNNTLSGELARRN-- 467
QY 465 NDPRKMDIDYNSVTVSNNEETVQVQAGNNNNENDSEMKVTLFNHPSCQQOANGNGSDQ 544
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 468 ---NNSIV--VRNTE-----DQ 479
QY 545 KIMNCGNRYNSAFSMALQDLIGIDVSGQHMLDESSKIGTHFSNTSSLYTSLSSREA 604
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 480 TALNA-----VEGGSNREV 494
QY 605 SPEKRGPSL--LFPMPMETKIV-NPIGNSVTSMLESPFVQMRPSPALISLHPVFAASMT 661
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 495 STPERLISFPAIPALPQVAVQKFGSGMNGMSPMTSNPAELK-TVALTLQMPVFAAMA 553
QY 662 DT 663
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 554 DS 555

```

## RESULT 7

US-10-059-911-22  
Sequence 22, Application US/10059911  
Publication No. US20030159180A1  
GENERAL INFORMATION:

APPLICANT: Fischer, Robert L.  
 APPLICANT: Mizukami, Yukiko  
 TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 FILE REFERENCE: 023070-090730US  
 CURRENT APPLICATION NUMBER: US/10/059,911  
 CURRENT FILING DATE: 2002-06-17  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 22  
 LENGTH: 555  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: structurally  
 OTHER INFORMATION: altered ANT protein ANTmutII, temperature sensitive  
 OTHER INFORMATION: mutant  
 US-10-059-911-22

Query Match 37.5%; Score 1313; DB 14; Length 555;  
 Best Local Similarity 43.1%; Pred. No. 7.9e-114;  
 Matches 311; Conservative 70; Mismatches 115; Indels 226; Gaps 22;

1 MKRINESNNTDGNHNLGFSLSPH-MKM-----EATSAATVPTTFM 43  
 1 MKSFCDNDNNHSTTNILGFSLSNMKGGREGREAIYSSSTSAATSSSVPLV 60  
 44 SPSGSHLSNFGMCGVGENGFHSPLTYMPLKSDGSLCILEALRSQTVVVPSPKLE 103  
 61 G---DNISNFGVCIQSANGGIIYSMSWPLRSDGSLCILEALRSQTVVVPSPKLE 117  
 104 DFLGATGTHGYSHERG--LSLDSIYNSQNAEAPNDLSQPP---RQGGHNSVQ 157  
 118 DFFG---THNNTSHKAMDLSLDSLFYNTTH---EENTTNFQEFFSPQTRNH--- 166  
 158 THPIYSLGACGLYAPLEETTKETHVSDCSLMPQTEGLKNWVAPTRFSTHQVLE 217  
 167 -----EERT--RNYGND-----PSLTHG----- 182  
 218 QQMNCGMNERNGVSLGSGGEL-QSLSMSPSGSSCVTAPS----- 261  
 183 -----GSPNGV-----YGEFQGSLSMSPSGSSCTGSHHQQNQNHQSNH 229  
 262 -----GTSVAVDAAKRGH---AKLGQKQPVHRSIDTRFGQTSQYRGVTR 304  
 230 QOISEALVETSVGFETTMAAASSTRGQEDVVVGQKQVHRKSIDTFGQTSQYRGVTR 289  
 305 HMTGRVYEAHLMDNSCKEGQTRKGRQVYLLGYDMEERKARAYDLAALKWGPSTHINS 364  
 290 HMTGRVYEAHLMDNSCKEGHRSKGRQVYLLGYDMEERKARAYDLAALKWGPSTHINS 349  
 365 IENVQVLEMKNSGRQVYVLAHLRKSQSGFSGASITRGVTRHQHGMQARIGRVAGNK 424  
 350 AENVQKIEEDMKNSGRQVYVLAHLRKSQSGFSGASITRGVTRHQHGMQARIGRVAGNK 409  
 425 DLYLTFSTQERAAAYDVAALIKFRGANAVTNFDISRYDVERIMASSNLAGEIARRKD 484  
 410 DLYLTFSTQERAAAYDVAALIKFRGANAVTNFDITRDVDRIMSSNLLSGELLARRN-- 467  
 485 NDRNKDIDYNSVTVSNNEETVOVQAGNNNNENDESEMKVLFNHPSCQQQOANGNSDQ 544  
 468 -----NNSIV--VANTE-----DQ 479  
 545 KIMNCGNYSASFMAQLDILIGDSVSGQHMLDESSKIGTHFSNTSLVTSLSRREA 604  
 480 TLAHA-----VYEGSNKEV 494  
 605 SPEKRGPSL--LFPMPMETKIV-NPIGTSVTSMPSPTVQRPSPALISHLVYPSWT 661  
 495 STPERLSPPAIFALPQVQKMFSGNMGMNPSMTSNPAELK-TVALTLQMPVFAAMA 553  
 662 DT 663

DB 554 DS 555

RESULT 8  
 US-10-059-911-5  
 Sequence 5; Application US/10059911  
 Publication No. US20030159180A1  
 GENERAL INFORMATION:  
 APPLICANT: Fischer, Robert L.  
 APPLICANT: Mizukami, Yukiko  
 TITLE OF INVENTION: The Regents of the University of California  
 FILE REFERENCE: 023070-090730US  
 CURRENT APPLICATION NUMBER: US/10/059,911  
 CURRENT FILING DATE: 2002-06-17  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 548  
 LENGTH: 548  
 TYPE: PRT  
 ORGANISM: Brassica napus  
 FEATURE:  
 OTHER INFORMATION: Canola AINTEGUMENTA (ANT)  
 US-10-059-911-5

Query Match 37.3%; Score 1307; DB 14; Length 548;  
 Best Local Similarity 43.6%; Pred. No. 2.8e-113;  
 Matches 309; Conservative 77; Mismatches 103; Indels 220; Gaps 25;

9 NTDDGNHNLGFSLSPHM-----KMEATSAATVPTTFYMSPEQ---SHLSNFGMXY 57  
 6 DNDDSITNLLGFSLSNMLKMGGEGALYSSSSSSSVARS--SVFQVLYGDNSSNYGVY 63  
 58 GVG-ENGNSFPLTYMPLKSDGSLCILEALKS-----QTVNVPVPSPLDPLGCA 109  
 64 GSNLAAREWYSQMSWVPLRSDGSLCILEALNRSSSHNNHHEQV---SSPXMEDF--- 115  
 110 TWGTHGYS-SHERG--LSLDSIYNSQNAEAPNDLSQPPRQGGHNSVQTHPIYSGLA 166  
 116 -FETHHNTSHKAMDLSLDSLFYNTTH---PNNNTNQE---FSPQTRNH--- 163  
 167 CHGLYAPLEETTKETHVSDCSLMPQTEGLKNWVAPTRFSTHQVLEQNMCGMN 226  
 164 -----EERT--RNYEND-----PGLTHG----- 179  
 227 ERNGVSLGSGGEL-QSLSMSPSGSSCVTA-----P 260  
 180 ---GGSFNVGVYGEFQGSLSMSPSGSSCTTASHHQQNQNHQOISEALVETSAGE 236  
 261 SGTDSVAVDAAKRGH-HAKLGQKQPVHRSIDTRFGQTSQYRGVTRHMTGRVYEAHLMDNS 319  
 237 TTMMAAAAKKRGQGVVVGQKQVHRKSIDTFGQTSQYRGVTRHMTGRVYEAHLMDNS 296  
 320 CKEGQTRKGRQVYLLGYDMEERKARAYDLAALKWGPSTHINSFISNVOVLEEMKQNS 379  
 297 FKKEGHSRGRQVYLLGYDMEERKARAYDLAALKWGPSTHINSFISNVOVLEEMKQNS 356  
 380 ROEYVLAHLRKSQSGFSGASITRGVTRHQHGMQARIGRVAGNKDLYLGTGTOEEAAE 439  
 357 ROEYVLAHLRKSQSGFSGASITRGVTRHQHGMQARIGRVAGNKDLYLGTGTOEEAAE 416  
 440 AYDVAALIKFRGANAVTNFDISRYDVERIMASSNLAGEIARRKNDPRNKDIDYNSIV 499  
 417 AYDVAALIKFRGANAVTNFDITRDVDRIMASSNLLSGELLARRNS-----NNSIV 465  
 500 TSYNNEE---TVQVQAGNNNNENDESEMKVLFNHPSCQQQOANGNSDQKIMNCGNYSNA 556  
 466 RNISDEBALITAVYNGSNKEV-----GSPERVLSF----- 496  
 557 FMAQLDILIGDSVSGQHMLDESSKIGTHFSNTSLVTSLSRREASPEKRGPSLLFP 616  
 497 -----PTIFA 501

QY 617 MPPMETKI--VNPISVTWMLSPFTVQMRPSPA--SLSHLPVFAWMDT 663  
DB 502 LPQVBRKMGAVVGV--NMSWTTNNADLK--TVSLTLQMPVFAWMDS 548

## RESULT 9

US-10-059-911-13  
Sequence 13, Application US/10059911  
Publication No. US20030159180A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
FILE REFERENCE: 023070-090730US  
CURRENT APPLICATION NUMBER: US/10/059,911  
CURRENT FILING DATE: 2002-06-17  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 13  
LENGTH: 452  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: structurally  
OTHER INFORMATION: altered ANT protein ANTDCI  
US-10-059-911-13

Query Match 35.2%; Score 1233.5; DB 14; Length 452;  
Best Local Similarity 52.6%; Pred. No. 1.6e-106;  
Matches 275; Conservative 42; Mismatches 79; Indels 127; Gaps 15;

QY 1 MKRINSNTDDGNHNMIGFSLSPH--MCM-----EATSATVPTTYM 43  
DB 1 MMSFCNDNDNNHNTNLLGFSLSNMWKKGGGREGALYSSTSSAATSSSVFPQLV 60  
QY 44 SPQSHLSNFGMCYGVGENGFHSPPLVNPPLKSDGSLCILEALKESQTOVWPTSSPKLE 103  
DB 61 G---DNTSNRGVGYGNNPGIYSHSVMPLRSDGSLCIMEALNNSHSHHQCDSPKVE 117  
QY 104 DPLGATMGTHEYGSHERG--LSLDSITYNSQNAEQPNRDLISQF---RQGHMSVQ 157  
DB 118 DFRG---THNNTSKKAMDLSLDFYNTTH--EPNTTNFQEFFSPQTRNH--- 166  
QY 158 THVYSGIACG--YQAPLEETKETHVSDCSLMPQWTEGLKMWAPTRFESTHQVLE 217  
DB 167 -----EET--RNYGND-----PSLTHG----- 182  
QY 218 QQANCGENRNGVSLGSGGEL--QSLSLNSPQSSGCVTAP----- 261  
DB 183 -----GSFNVGV-----YGFQOQSLSLMSPGSSCTIGSHHQNQONQONH 229  
QY 262 -----GTDVAVDAKKGH---AKLGQKOPVHRKSIDTFGORTSQYRGVTR 304  
DB 230 QQISEALVETSVGFETTTMAAKKGQEDVYVVGQKQIVHRKSIDTFGORTSQYRGVTR 289  
QY 305 HRMTGRYEAHLMDNSCKKSGQTRKGRQVYIGYDMEKARAYDLAALKYWGSPSTHTNS 364  
DB 290 HRMTGRYEAHLMDNSCKKSGQTRKGRQVYIGYDMEKARAYDLAALKYWGSPSTHTNS 349  
QY 365 IENVOLEEMKMSRQEVYVAHLRKSSGFSRGASITYRGVTRHGHGRMQRARIGRYAGK 424  
DB 350 AENQKIEDKMKNTROEVYVAHLRKSSGFSRGASITYRGVTRHGHGRMQRARIGRYAGK 409  
QY 425 DLYLGTFTSTOEAAAYDAVAALFKRGANAVTNFDISRYVERI 467  
DB 410 DLYLGTFTSTOEAAAYDAVAALFKRGANAVTNFDITRYDRI 452

RESULT 10  
US-10-059-911-9  
Sequence 9, Application US/10059911  
Publication No. US20030159180A1

GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
FILE REFERENCE: 023070-090730US  
CURRENT APPLICATION NUMBER: US/10/059,911  
CURRENT FILING DATE: 2002-06-17  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 473  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: structurally  
OTHER INFORMATION: altered ANT protein ANTDCI  
US-10-059-911-9

Query Match 35.0%; Score 1224; DB 14; Length 473;  
Best Local Similarity 44.9%; Pred. No. 1.3e-105;  
Matches 286; Conservative 56; Mismatches 89; Indels 206; Gaps 19;

QY 69 LTYWPLKSDGSLCILEALKESQTOVWPTSSPKLEDFLGATMGTHEYGSHERG--LSLD 126  
DB 1 MSVWPLRSDGSLCIMEALNNSHSHHODSSPKVEDFFG---THNNTSKKAMDLSLD 56  
QY 127 SIYNSQNAEQPNRDLISQF---RQGHMSVQTHPYYSGLACGLVQAPLEETKE 182  
DB 57 SLFNTH---EPNTTNFQEFFSPQTRNH-----EET--R 89  
QY 183 THVSDCSLMPQWTEGLKMWAPTRFESTHQVLEQONCGMGENRNGVSLGSGGEL-- 241  
DB 90 NYGND-----PSLTHG-----GSFNVGV-----YGFQ 112  
QY 242 QSLSLNSPQSSGCVTAP-----GTDVAVDAK 272  
DB 113 QSLSLNSPQSSGCVTAP-----GTDVAVDAK 172  
QY 273 RGH---AKLGQKOPVHRKSIDTFGORTSQYRGVTRHMTGRYEAHLMDNSCKKSGQTRG 329  
DB 173 RGHSDVYVVGQKQIVHRKSIDTFGORTSQYRGVTRHMTGRYEAHLMDNSCKKSGQTRG 232  
QY 330 RQVTLGGYDMEKARAYDLAALKYWGSPSTHTNSIENVOLEEMKMSRQEVYVAHLR 389  
DB 233 RQVTLGGYDMEKARAYDLAALKYWGSPSTHTNSIENVOLEEMKMSRQEVYVAHLR 292  
QY 390 KSGFSRGASITYRGVTRHGHGRMQRARIGRYAGKDIYLGFTSTOEAAAYDAVAALFKR 449  
DB 293 KSGFSRGASITYRGVTRHGHGRMQRARIGRYAGKDIYLGFTSTOEAAAYDAVAALFKR 352  
QY 450 GANAATVNDISRYVERIMASSNLLAGELARKKNDPRKQIDYKSVTSTVNNETVQ 509  
DB 353 GTNATVNDITRYDRIYVDRIMSSNTLLSGELARRN-----NNSIV--YANIE-- 395  
QY 510 VQAGNNNENDESEKQVLPHPSSQOQANGSDOKIMCNGYRNSAFMALDLDIGIDS 569  
DB 396 -----DQTLN----- 402  
QY 570 VGSQOHMLDESSKIGTHFGNTSLVTSLSRRASEKRGPSL--LFPMPMETKI--N 626  
DB 403 -----VEGGSNKEVSTPERLSPALFALPQVQKMGFS 437  
QY 627 PIGTSYTWMLSPFTVQMRPSPA--SLSHLPVFAWMDT 663  
DB 438 NMGGMSFWTSNPAALIK--YVALLTLQMPVFAWMDS 473

RESULT 11  
US-10-024-632-11  
Sequence 11, Application US/10024632  
Publication No. US20020170093A1  
GENERAL INFORMATION:

/ APPLICANT: Monsanto Technology LLC  
 / APPLICANT: He, Steve S.  
 / APPLICANT: Dotson, Stanton B.  
 / TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION  
 / TITLE OF INVENTION: GROWTH AND USES THEREOF  
 / FILE REFERENCE: 38-21(5183)B  
 / CURRENT APPLICATION NUMBER: US/10/024,632  
 / CURRENT FILING DATE: 2001-12-19  
 / PRIOR APPLICATION NUMBER: US 60/257,896  
 / PRIOR FILING DATE: 2000-12-21  
 / NUMBER OF SEQ ID NOS: 33  
 / SOFTWARE: PatentIn version 3.1  
 / SEQ ID NO 11  
 / LENGTH: 585  
 / TYPE: PRT  
 / ORGANISM: Gossypium hirsutum  
 / US-10-024-632-11

Query Match 33.7%; Score 1180.5; DB 13; Length 585;  
 Best Local Similarity 41.7%; Pred. No. 2,2e-101; Indels 163; Gaps 25;  
 Matches 290; Conservative 79; Mismatches 164;

QY 17 NMWGLSLSPHMKMEATSAATVPTTFTYSPSOSHLNFGMCYGVGENGNFHSPLTVMPLKS 76  
 DB 3 NMWGLSLTPTDLRIDES-----FGREDHGFPSS---VMDLRS 35  
 QY 77 DGSICLLEALKRSQT-----QVWVPTSSPKLEDFLG---GATMGTHEY-GS 118  
 DB 36 DSSLGVDPFRSSSTIADBDREYENGIGSATNDEGPDLPDLCYSNPSQETKAYCGT 95  
 QY 119 HER-----GLSLDSIYNSQNAARQPN-----BDLSQP--FRQGHMS 155  
 DB 96 HENONTVPSPRTINNVAVAPNYSSSGDAEAENFTNPSFPIOTYNNYENPQTLMAQHS 155  
 QY 156 VQHPHYSGLAGCHLYQAPLEETTKETHVSPDSSIMPMQTEGLKNWVAPRREKSTHOV 215  
 DB 156 QCCDPNPNHNSGVAHVHPESATS-----VSGRKSMLRQI--PPGSK-- 197  
 QY 216 LEOQNMNGNENRNGVSLGSGELQSLSLMSBPSQSCVTPASGTDVYA---VDAK 271  
 DB 198 -----ASGNETNNF-----NFOALSLTMSB-----TSRNGFPALAPLEVVDNR 236  
 QY 272 KSGHAKLGOKOPVHRKSIDTFGORTSOYRGVTRHRTGRYEAHLMDNSCKKSGOTRKG 331  
 DB 237 KRPVGNLTRRESVPRKSIDTFGORTSOYRGVTRHRTGRYEAHLMDNSCKKSGOTRKG 236  
 QY 332 VYLGVDMEKARAYDLAALKYWPSTHINFSIENYQVLEEMKNMSROEYVAHLRR 391  
 DB 297 VYLGVDMEKAKAYDLAALKYWPSTHINFSIENYQVLEEMKNMTRQCFVAHLRR 356  
 QY 392 SSGSRASISYRGVTRHOGHGRQWARIIGVYAGNKDLYLGTFTQEEAAEYVAUAIKFR 451  
 DB 357 SSGSRASISYRGVTRHOGHGRQWARIIGVYAGNKDLYLGTFTQEEAAEYVAUAIKFR 416  
 QY 452 NAYTNFDIRYVERIMASSNLLAGELARRKNDPRNKDIDYKSVVTSYNNETVQV 511  
 DB 417 SAVTNFDIRYVERIKSSSTLIGELAKRSPKDTASIAPEVDN--SCASSAPQPLAIP 475  
 QY 512 AGNNNNE--NDSEKMYLFWHPSSQOO--QANGNSDCKINNCNYYNSAFVMAJODLIGI-- 567  
 DB 476 SGEASBELADMTWTA---NSDEQOQHOSNTNNDASLANSS--RNSNPSQSPKSTGLAS 531  
 QY 568 DVSQSGQNMLEDSSKIGTHPSNTSLVTSJSSSEASPEKRGPSLLFPMPMETKLVN 627  
 DB 532 DKRGIG---GYSYHG--YFSLKSKYEDGNSSEIDNSNENR----- 567  
 QY 628 IGTSVTSMLPSPVQWRPSPALSLH--LPVFASTMD 662  
 DB 568 LGN-----LGLVHKIPMPALMNE 585

RESULT 12  
 US-10-059-911-15

/ Sequence 15, Application US/10059911  
 / Publication No. US20030159180A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Fischer, Robert L.  
 / APPLICANT: Mizukami, Yukiko  
 / TITLE OF INVENTION: The Regents of the University of California  
 / TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 / FILE REFERENCE: 023070-090730US  
 / CURRENT APPLICATION NUMBER: US/10/059,911  
 / CURRENT FILING DATE: 2002-06-17  
 / NUMBER OF SEQ ID NOS: 42  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 15  
 / LENGTH: 370  
 / TYPE: PRT  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / OTHER INFORMATION: Description of Artificial Sequence: structurally  
 / OTHER INFORMATION: altered ANT protein ANTDMCI  
 / US-10-059-911-15

Query Match 32.3%; Score 1129.5; DB 14; Length 370;  
 Best Local Similarity 56.4%; Pred. No. 6.5e-97;  
 Matches 247; Conservative 29; Mismatches 55; Indels 107; Gaps 12;

QY 69 LTVMPKSDGSLCILEALKRSQIQVWPTSSPKLEDFLGATMGTHEYSGHERG--LSLD 126  
 DB 1 MSVMPLRBDSLCLEALNRRSSHHNODSPKVEDFG---THNNNTSHKEMDLSLD 56  
 QY 127 SIYNSQNAARQPNRDLISQPF---RQGHMSVQTHPHYSGLAGCHLYQAPLEETTK 182  
 DB 57 SLFVNTTH---EPNTTNFOCFEFSFPQTRNH-----EEET--R 89  
 QY 183 TVSPDSSIMPMQTEGLKNWVAPRREKSTHOVLEQNMNGNENRNGVSLGSGCEL-- 241  
 DB 90 NYGND-----PSLTHG-----GSFNVGV-----YGFQ 112  
 QY 242 QSLSLMSBPSQSCVTPAS-----GTDSVAVDARK 272  
 DB 113 QSLSLMSBPSQSCSITGSHHCONONHOSQNHQOISALVETSYGFETTTMAAKKK 172  
 QY 273 RGH---AALGOKOPVHRKSIDTFGORTSOYRGVTRHRTGRYEAHLMDNSCKKSGOTRKG 329  
 DB 173 RQGEVYVYVGGQKQYHRKSIDTFGORTSOYRGVTRHRTGRYEAHLMDNSCKKSGOTRKG 232  
 QY 330 RQVYLGVDMEKARAYDLAALKYWPSTHINFSIENYQVLEEMKNMSROEYVAHLRR 389  
 DB 233 RQVYLGVDMEKARAYDLAALKYWPSTHINFSIENYQVLEEMKNMTRQCFVAHLRR 292  
 QY 390 KSSGSRASISYRGVTRHOGHGRQWARIIGVYAGNKDLYLGTFTQEEAAEYVAUAIKFR 449  
 DB 293 KSSGSRASISYRGVTRHOGHGRQWARIIGVYAGNKDLYLGTFTQEEAAEYVAUAIKFR 352  
 QY 450 GNAVTNFDISRYVERI 467  
 DB 353 GNAVTNFDITRYVDRI 370

RESULT 13  
 US-10-059-911-10  
 / Sequence 10, Application US/10059911  
 / Publication No. US20030159180A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Fischer, Robert L.  
 / APPLICANT: Mizukami, Yukiko  
 / TITLE OF INVENTION: The Regents of the University of California  
 / TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
 / FILE REFERENCE: 023070-090730US  
 / CURRENT APPLICATION NUMBER: US/10/059,911  
 / CURRENT FILING DATE: 2002-06-17  
 / NUMBER OF SEQ ID NOS: 42  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 10

LENGTH: 308  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: structurally  
OTHER INFORMATION: altered ANT protein ANTND2  
US-10-059-911-10

Query Match 29.8%; Score 1042; DB 14; Length 308;  
Best Local Similarity 54.0%; Pred. No. 7,7e-89;  
Matches 218; Conservative 36; Mismatches 48; Indels 102; Gaps 8;

QY 266 VAVDAKKGH--AKLGKQKPYHRSIDTFQGRISQYRGVTRHRTGTYEAHLWDSCKK 322  
DB 1 MAIAKKKQDEVVVVGKQIVHRKSIDTFQGRISQYRGVTRHRTGTYEAHLWDSCKK 60  
QY 323 ECGTRKGRQVYIGGIDMEKARAYDLAKTKWGSTHINSIENYQVLEEMKMSROE 382  
DB 61 EGHSRKGRQVYIGGIDMEKARAYDLAKTKWGSTHINSIENYQVLEEMKMSROE 120  
QY 383 YVAHLRRKSGFSGASIRGVTRHHRGWRQARIGRVAGNKDYLGTFTQEEAAERYD 442  
DB 121 YVAHLRRKSGFSGASIRGVTRHHRGWRQARIGRVAGNKDYLGTFTQEEAAERYD 180  
QY 443 VAAIKRGANAVTNPDISRYVERIMASNLLAGELARRKNDPRNDIDYNSVTVSY 502  
DB 181 VAAIKRGANAVTNPDITRYDVRIMSSNTLLSGELARRN--NNSIV--V 226  
QY 503 NNEETVQVQAGNNNNNDESEWKKVLFNHPSCQQAANGSDOKIMCGYNSAFSMAIQ 562  
DB 227 RUTE-----DGTALNA----- 237  
QY 563 DLIGDSVSGQHMLDESSKIGTHFSNTSLVTSLSSEKRSPEKRGSL--LFPMPM 620  
DB 238 -----VVEGGSNKEVSTPERLLSPFAIFALPOV 265  
QY 621 ETKIV-NPIGTSVSWLSPYVQMRPSPAISLSHLFPFASWTD 663  
DB 266 NQPMFGSNMGANSPWTSNFMELK-TVALTLFQMFVFAWADS 308

## RESULT 14

US-10-059-911-11  
Sequence 11, Application US/10059911  
Publication No. US20030159180A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Mizukami, Yukiko  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Methods for Altering Organ Mass in Plants  
FILE REFERENCE: 023070-090730US  
CURRENT APPLICATION NUMBER: US/10/059,911  
CURRENT FILING DATE: 2002-06-17  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 277  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: structurally  
OTHER INFORMATION: altered ANT protein ANTND3  
US-10-059-911-11

Query Match 27.3%; Score 955.5; DB 14; Length 277;  
Best Local Similarity 53.2%; Pred. No. 8,3e-81;  
Matches 198; Conservative 34; Mismatches 41; Indels 99; Gaps 7;

QY 295 RTSQYRGVTRHRTGTYEAHLWDSCKKEGTRKGRQVYIGGIDMEKARAYDLAKTK 354  
DB 2 RTSQYRGVTRHRTGTYEAHLWDSCKKEGTRKGRQVYIGGIDMEKARAYDLAKTK 61  
QY 355 WGPSTHINSIENYQVLEEMKMSROEYVAHLRRKSGFSGASIRGVTRHHRGWRQ 414

## RESULT 15

US-10-374-780A-334  
Sequence 334, Application US/10374780A  
Publication No. US20040019927A1

DB 62 WGPSTHINSIENYQVLEEMKMSROEYVAHLRRKSGFSGASIRGVTRHHRGWRQ 121  
QY 415 ARIGRVAGNKDYLGTFTQEEAAERYDVAIAIKFRGANAVTNFDISRYVERIMASNL 474  
DB 122 ARIGRVAGNKDYLGTFTQEEAAERYDVAIAIKFRGANAVTNFDITRYDVRIMSSNTLL 181  
QY 475 AGEIARRKNDPRNKDIDYNSKVTSVNNETVQVQAGNNNNDESEWKKVLFNHPSCQ 534  
DB 182 SGELARRN-----NNSIV--VANTE----- 199  
QY 535 QQAANGSDOKIMCGYNSAFSMAIQDLIGDSVSGQHMLDESSKIGTHFSNTSL 594  
DB 200 -----DGTALNA----- 206  
QY 595 VTSLSRRASPEKRGPSL--LFPMPMETKIV-NEIGTSVSWLSPYVQMRPSPAISL 651  
DB 207 VVEGGSNKEVSTPERLLSPFAIFALPOVQKFGSNMGANSPWTSNFMELK-TVALTL 265  
QY 652 SHLPVPSWTD 663  
DB 266 PQMPVFAWADS 277

GENERAL INFORMATION:  
APPLICANT: Sherman, Bradley K  
APPLICANT: Riechmann, Jose Luis  
APPLICANT: Jiang, Cai-Zhong  
APPLICANT: Heard, Jacqueline E  
APPLICANT: Haake, Volker  
APPLICANT: Creelman, Robert A  
APPLICANT: Raccilite, Oliver  
APPLICANT: Adam, Luc J  
APPLICANT: Reuber, T. Lynne  
APPLICANT: Keddie, James  
APPLICANT: Brown, Pierre E  
APPLICANT: Piglium, Marsha L  
APPLICANT: Dubell III, Arnold T  
APPLICANT: Phineda, Omaira  
APPLICANT: Yu, Guo-Liang  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
FILE REFERENCE: MBI-0047 CIP  
CURRENT APPLICATION NUMBER: US/10/374,780A  
CURRENT FILING DATE: 2003-02-25  
PRIOR APPLICATION NUMBER: 09/837,944  
PRIOR FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: 60/310,847  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/934,455  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/336,049  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/338,692  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: 10/171,468  
PRIOR FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 10/225,066  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,067  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,068  
PRIOR FILING DATE: 2002-08-09  
NUMBER OF SEQ ID NOS: 2906  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 334  
LENGTH: 574  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:

OTHER INFORMATION: G1793 (conserved domain in AA coordinates:179-255, 281-349)  
US-10-374-780A-334

Query Match 27.1% Score 948; DB 15; Length 574;  
Best Local Similarity 35.9% Pred No 1,4e-79;  
Matches 257; Conservative 86; Mismatches 164; Indels 208; Gaps 26;

```
QY 14 NNNNTLIGPSLSPHMKMEATSAATVPTTFMSPSGHLSN-----FGMCYGVGENGFHS 67
DB 2 NSNNMLGPELSPN-----NSSLPPEHYVLGLVSDHMDNPFQTEMNMINPHGGGD--- 52
QY 68 PLTVMLKSDGLICILALKRSGQTQVMPTSPKLEDFLGATMGTHGYSGHERGLSDS 127
DB 53 -----EGG-----EVPKADPLG--VSKPDENGNHLVAYNDS 83
QY 128 IYNSQNAEAQPNRDLISQPFROGMSVQHPYYSGLACHGLYQAPLEETKETHVSD 187
DB 84 DYFFHTNS-LMP-----SYQSN-----DVVVA 105
QY 188 GSLMPQMTGELKNWVAPTRFPTTHQOYLEQQMNGMERNNGVSLSGGGLSLSL 247
DB 106 CSNTPTNN-----SYHELQESNN-----LSLTL 132
QY 248 MSPSGSSCVTAPSGTDSVAVDAKKRGAHLGQKQPVHRKSIDTFGQRTSQYRGVTRRW 307
DB 133 MGTAGNNVVDKASSETTGDNASGALAVETATP--RRALDTFGORTSIYRGVTRRW 190
QY 308 TGRYEALHMDNSCKEGGTRKGRQYLLGGYMEEKARAYDLAALKYMGPSHTINFSIEN 367
DB 191 TGRYEALHMDNSCKEGGTRKGRQYLLGGYMEEKARAYDLAALKYMGPSHTINFSIEN 250
QY 368 YQVQLEEMKNGRQEVVAHLRRKSSGFSRGASIRGVTRHHQGRWQARIQVAGNKDLY 427
DB 251 YEKVEYEMKMTROEFVAALIRKSSGFSRGASIRGVTRHHQGRWQARIQVAGNKDLY 310
QY 428 LGFTSTQEEAEAYDVAATKEFGANAVTNFDISRVDVERIMASSNL-LAGELARKKD-- 484
DB 311 LGFTSTEEBAAYDIAIKFRGLNAVTFEIRNRYDAKALLESSTLPIGGGAARLKEAQ 370
QY 485 --NDPNKCID-----YKNSV-----VTSVNNEET 507
DB 371 ALESSRKREAEMLALGSSFOYGGSGSTGSTRLOLPYPLSLIQPLEPFLSLQND- 429
QY 508 VOYQAGNNNNENDSEMKWLFNHPGQ-----QQANGGSDPKIMNGNY 552
DB 430 --ISHNNNNADDS---SFNHSYITQTLHQQTNNYLOQSSON--SQCLYNAYLH 481
QY 553 RNSAFSMLQDLIGIDSVSGGQHNLDESSKIGTHFSNTSLVTSL-----SSSRKSPPE 607
DB 482 SNPAL--LHGLVSTSIYDNNNNN---GSSGSY--NTAFLGNHIGIGSSSTVGSTE 532
QY 608 KRGPILLFPMPEMKIYVPIGSTVSWLPSPTQMRSPALISHLPVFASWTD 662
DB 533 E-----PPTYKTIDYDMPSDGTGCGYSGWTSBSVO--GSNPG-----GVFTYMNNE 574
```

Search completed: March 9, 2004, 10:52:46  
Job time: 40 secs